

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 06:17:00 ; Search time 11147.1 Seconds
(without alignments)
11447.105 Million cell updates/sec

Title: US-10-669-693-1

Perfect score: 2944

Sequence: 1 ttctctgttgcttactc.....aaaaaaaaaaaaaaaaaaaaa 2944

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2622.4	89.1	2626	9	AY358603	AY358603 Homo sapi
2	2621.4	89.0	2623	6	AX463930	AX463930 Sequence
3	2615.8	88.9	2619	9	BC033691	BC033691 Homo sapi
4	2612	88.7	2620	6	AX327331	AX327331 Sequence
5	2599.2	88.3	2604	6	AX430320	AX430320 Sequence
6	2596	88.2	2604	6	AX552230	AX552230 Sequence
7	2399.4	81.5	2513	6	AX876954	AX876954 Sequence
8	2399.4	81.5	2513	6	BD156406	BD156406 Primer fo
9	2399.4	81.5	2513	9	AK027605	AK027605 Homo sapi
10	2233.2	75.5	2261	6	AX780433	AX780433 Sequence
11	2233.2	75.5	2261	9	AF335278	AF335278 Homo sapi
12	1717.4	58.3	2114	6	AX552231	AX552231 Sequence
13	1513.4	51.4	1515	6	AX591678	AX591678 Sequence
14	1265.2	43.0	129402	9	AC011510	AC011510 Homo sapi
15	1263.6	42.9	173251	2	AC058798	AC058798 Homo sapi
16	885.6	30.1	6192	6	AX251439	AX251439 Sequence
17	658.4	22.4	2081	10	BC034202	BC034202 Mus muscu
18	613.2	20.8	6192	6	AX251440	AX251440 Sequence
19	601.2	20.4	812	6	AX381127	AX381127 Sequence
20	560.8	19.0	567	6	AR380149	AR380149 Sequence
21	556.8	18.9	562	6	AX779032	AX779032 Sequence
22	496.8	16.9	2065	4	RABCYCP2C	M20857 Rabbit cyto
23	495.2	16.8	2026	4	RABP450BX	L10912 Rabbit cyto
24	495.2	16.8	2028	4	S64259	S64259 cytochrome
25	490.4	16.7	2045	4	RABCYCP2B	M20856 Rabbit cyto
26	490.2	16.7	548	6	AX868269	AX868269 Sequence
27	490.2	16.7	548	6	BD148331	BD148331 Primer fo
28	489.2	16.6	1641	4	RABIIA10	L10236 Oryctolagus
29	488.6	16.6	1781	4	AB052255	AB052255 Sus scrof
30	485.6	16.5	2079	4	RABCYCF2A	M20855 Rabbit cyto
31	485.2	16.5	537	6	AX873328	AX873328 Sequence
32	485.2	16.5	537	6	BD153390	BD153390 Primer fo
33	482.8	16.4	2484	4	RABIIA11	L10237 Oryctolagus
34	482.4	16.4	1717	10	D86952	D86952 Mesocricetu
35	482.4	16.4	1802	4	RABP450CA	M18820 Rabbit cyto
36	481.2	16.3	2743	4	AB052256	AB052256 Sus scrof
37	475.2	16.1	1485	6	E10626	E10626 Human cDNA
38	475.2	16.1	1485	6	E10859	E10859 cDNA encodi
39	475.2	16.1	1485	6	E10860	E10860 cDNA encodi
40	475.2	16.1	1485	6	AR399327	AR399327 Sequence
41	475.2	16.1	1485	6	AR399328	AR399328 Sequence
42	475.2	16.1	1748	6	AX409692	AX409692 Sequence
43	475.2	16.1	1748	9	HUMCPIIA3A	M33318 Human cytoc
44	474.8	16.1	1485	9	AF209774	AF209774 Homo sapi
45	474.2	16.1	2004	6	AX477702	AX477702 Sequence

ALIGNMENTS

RESULT 1
AY358603
LOCUS
DEFINITION Homo sapiens clone DNA76395 Cytochrome p450 (UNQ891) mRNA, complete cds.
ACCESSION AY358603
VERSION AY358603.1 GI:37182327
KEYWORDS FLI_CDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2626)
AUTHORS Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
Chen,J., Chow,B., Chui,C., Crowley,C., Deuel,B.,

QY	1763	CCTTCCAGCTGCAAGTGTGGTGGGTGCCCAGGACGGTGCCTCCAGCCCTCACTCCACACGACGACGACGATGAAGGA	1822
Db	1503	CCTTCCAGCTGCAAGTGTGGTGGGTGCCCAGGACGGTGCCTCCACACGACGACGACGATGAAGGA	1562
QY	1823	AGGCAACTTTGGAAGTGTGGTGGGTGCCCAGGACGGTGCCTCCAGCCCTCAACAGTGGCGATGG	1882
Db	1563	AGGCAACTTTGGAAGTGTGGTGGGTGCCCAGGACGGTGCCTCCAGCCCTCAACAGTGGCGATGG	1622
QY	1883	ACAGGTTTAATGTCTCCAGAGTGTACACTGCAGAGCAGCCACATTTACACGCCTGCAGTTG	1942
Db	1623	ACAGGTTTAATGTCTCCAGAGTGTACACTGCAGAGCAGCCACATTTACACGCCTGCAGTTG	1682
QY	1943	TTTTCCGGAGTCTGTCCACAGGCCACACGGTCACTTTAGCTGGGTAGCTTTTCCACA	2002
Db	1683	TTTTCCGGAGTCTGTCCACAGGCCACACGGTCACTTTAGCTGGGTAGCTTTTCCACA	1742
QY	2003	ACCGCACCCATACAACTTACAAGGGCCACAAAGCAAATGCTGGGTAGCTTTTCCACA	2062
Db	1743	ACCGCACCCATACAACTTACAAGGGCCACAAAGCAAATGCTGGGTAGCTTTTCCACA	1802
QY	2063	GACATAAATATAGTCCATCTGCAATATCAAGACATAGCCAGGTAACCCACCACTCCCC	2122
Db	1803	GACATAAATATAGTCCATCTGCAATATCAAGACATAGCCAGGTAACCCACCACTCCCC	1862
QY	2123	TGATCTGCAGCCACACAGTGGGAGTCTGGCTGTCACTTTCACAGGCCACAGAAAAGGCC	2182
Db	1863	TGATCTGCAGCCACACAGTGGGAGTCTGGCTGTCACTTTCACAGGCCACAGAAAAGGCC	1922
QY	2183	ACACATGTTTACAGCTCACACGCCCTCTCCATTATCGAACTTCTCAGTGTCCCTGTGCC	2242
Db	1923	ACACATGTTTACAGCTCACACGCCCTCTCCATTATCGAACTTCTCAGTGTCCCTGTGCC	1982
QY	2243	TGTTGCTCCGACAGGAAACAGATGCCCCCTCCGGGTTCATGCCACCCACAGAGCTGTGC	2302
Db	1983	TGTTGCTCCGACAGGAAACAGATGCCCCCTCCGGGTTCATGCCACCCACAGAGCTGTGC	2042
QY	2303	CTGTCTATGGCCCCCACTCATGTCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGAC	2362
Db	2043	CTGTCTATGGCCCCCACTCATGTCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGAC	2102
QY	2363	CACCGATGTCCACACACCCCCAACCTTGTCCACACAGCTACCCAGTACGACATCGTC	2422
Db	2103	CACCGATGTCCACACACCCCCAACCTTGTCCACACAGCTACCCAGTACGACATCGTC	2162
QY	2423	CTGGCTCCCCAGAGTATCTTCCACTGAGACACGCGCGCCCCCAGAGGCACAGTCCCCA	2482
Db	2163	CTGGCTCCCCAGAGTATCTTCCACTGAGACACGCGCGCCCCCAGAGGCACAGTCCCCA	2222
QY	2483	GCCACCTCTGCAACTGACGGCCTCAGTCACCCCTTTTAAAGACCCCTGATTTACCAAT	2542
Db	2223	GCCACCTCTGCAACTGACGGCCTCAGTCACCCCTTTTAAAGACCCCTGATTTACCAAT	2282
QY	2543	GCAAAACACATCTGGGTCTGGGATTATGACAGAGACTTTGGACATACAGGACCCCTCAGA	2602
Db	2283	GCAAAACACATCTGGGTCTGGGATTATGACAGAGACTTTGGACATACAGGACCCCTCAGA	2342
QY	2603	CCGAGGAAACACTGCCCCAACCCCCAACAGTGTCTATGTAAACACGCTGGAAAGCGGCCCC	2662
Db	2343	CCGAGGAAACACTGCCCCAACCCCCAACAGTGTCTATGTAAACACGCTGGAAAGCGGCCCC	2402
QY	2663	TGCTGCCCTTCCACACACATACACTGTATCTACAGCCCTCTGTTCGGCGTTCAGA	2722
Db	2403	TGCTGCCCTTCCACACACATACACTGTATCTACAGCCCTCTGTTCGGCGTTCAGA	2462
QY	2723	GTCCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAAATTCAC	2782
Db	2463	GTCCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAAATTCAC	2522
QY	2783	CCTGTCCAGGAGTGAGCCGGATCTGACGTTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT	2842
Db	2523	CCTGTCCAGGAGTGAGCCGGATCTGACGTTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT	2582

QY	2843	TAAAGTTTGTCTGGCCTTAGCTAAAAA	linear	PAT 16-JUL-2002
Db	2583	TAAAGTTTGTCTGGCCTTAGCTAAAAA		
RESULT 2				
AX463930				
LOCUS	AX463930	2623 bp	DNA	
DEFINITION	Sequence 63 from Patent WO0140466.			
ACCESSION	AX463930			
VERSION	AX463930.1	GI:21898982		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E., Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L., Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K., Wood,W.L. and Zhang,Z.			
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding same			
JOURNAL	Patent: WO 0140466-A 53 07-JUN-2001;			
FEATURES	Genotech Inc. (US)			
source	Location/Qualifiers			
	1..2623			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
	Query Match	89.0%; Score 2621.4; DB 6; Length 2623;		
	Best Local Similarity	100.0%; Pred. No. 0;		
	Matches 2622; Conservative	0; Mismatches	1; Indels	0; Gaps
QY	264	GCGGAGCGCTGGGAGGAGGAGCGACCTGCCAGATGGAGCGACGCACCTG	323	
Db	1	GCGGAGCGCTGGGAGGAGGAGCGACCTGCCAGATGGAGCGACGCACCTG	60	
QY	324	GCGCTGCTGCTGGCGCTGGCGTGCTCTCTGCTGTGACGCTGGCGCTCTCGGGACCAG	383	
Db	61	GCGCTGCTGCTGGCGCTGGCGTGCTCTCTGCTGTGACGCTGGCGCTCTCGGGACCAG	120	
QY	384	GGCCCGAGGCGCACTGCCCGCCCGGGCCACGGCGCTAACACTGCTGGGAAA	443	
Db	121	GGCCCGAGGCGCACTGCCCGCCCGGGCCACGGCGCTAACACTGCTGGGAAA	180	
QY	444	GTTACGGCGCGGGCGCTGTATTACGGGCTCATGCGGCTGAGTAAGAAGTACGGACCGGT	503	
Db	181	GTTACGGCGCGGGCGCTGTATTACGGGCTCATGCGGCTGAGTAAGAAGTACGGACCGGT	240	
QY	504	GTTACCATCTACCTGGGACCTGGCGCTGTGTGTCTGTTGGCAGGAGCTGT	563	
Db	241	GTTACCATCTACCTGGGACCTGGCGCTGTGTGTCTGTTGGCAGGAGCTGT	300	
QY	564	GCGGAGGCGCTGGGAGGTCAGCTGAGAGTTACGGSCGGGAAACGATAGCATGCT	623	
Db	301	GCGGAGGCGCTGGGAGGTCAGCTGAGAGTTACGGSCGGGAAACGATAGCATGCT	360	
QY	624	GGAAGGACTTTTGATGGCCATGGGGTTTTCTTCTCAAACGGGAGCGGTGGAGGCAGCT	683	
Db	361	GGAAGGACTTTTGATGGCCATGGGGTTTTCTTCTCAAACGGGAGCGGTGGAGGCAGCT	420	
QY	684	GAGGAAGTTTACCATGCTGCTCTCGGGAACCTGGGCAATGGGGAAGCGAGGCGAGGA	743	
Db	421	GAGGAAGTTTACCATGCTGCTCTCGGGAACCTGGGCAATGGGGAAGCGAGGCGAGGA	480	
QY	744	GCTGATCCAGGCGGAGGCCGGTGTCGTGGAGACATTCAGGGGACAGAGGACGCC	803	
Db	481	GCTGATCCAGGCGGAGGCCGGTGTCGTGGAGACATTCAGGGGACAGAGGACGCC	540	
QY	804	ATTGATCCCTCCCTGCTGCTGGCCAGGCACCTTCCAACGTAGTCTGCTCCCTCTCTT	863	

[illegible]

BC033691.1	GI:21707057
MG	
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REMARK	
COMMENT	
FEATURES	
source	
gene	

Db 601 ATAAGGAGTTCAGGCGGTGGTCCGGGACAGTGGTGTACCCCTGCTGGAGTCAGCTCCC 660
 Qy 947 AGGGGGGTACAGCTACAGAGATGTTCTCTGTGTTCTGTGGGCGCCCTGCCAGGCGCCCA 1006
 Db 661 AGGGGGGTACAGCTACAGAGATGTTCTCTGTGTTCTGTGGGCGCCCTGCCAGGCGCCCA 720
 Qy 1007 AGCAGCTCTCCACACAGTCAGCACCTTGGCTGGCTTTCACAGTCGGGACAGGTGCAGCAGC 1066
 Db 721 AGCAGCTCTCCACACAGTCAGCACCTTGGCTGGCTTTCACAGTCGGGACAGGTGCAGCAGC 780
 Qy 1067 ACCAGGGAACTGGATGCTTCGGGCGCCGCGACAGTGTGCGATGCTTTCCTGCTGA 1126
 Db 781 ACCAGGGAACTGGATGCTTCGGGCGCCGCGACAGTGTGCGATGCTTTCCTGCTGA 840
 Qy 1127 AGATGGCACAGGAGGAACAAACCCAGGACAGAGATTACCAACAGAAACATGCTGATGA 1186
 Db 841 AGATGGCACAGGAGGAACAAACCCAGGACAGAGATTACCAACAGAAACATGCTGATGA 900
 Qy 1187 CAGTCATTTATTTGCTGTTGCTGGGACGATGACGCTCAGCACACCGTGGCTATACCC 1246
 Db 901 CAGTCATTTATTTGCTGTTGCTGGGACGATGACGCTCAGCACACCGTGGCTATACCC 960
 Qy 1247 TCCTGCTCTGATGAATAACCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGGG 1306
 Db 961 TCCTGCTCTGATGAATAACCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGGG 1020
 Qy 1307 AGCTGGGGGTGGCCAGGACCAAGCCTAGGGACCGTACCGCCTCCCTTACACGACG 1366
 Db 1021 AGCTGGGGGTGGCCAGGACCAAGCCTAGGGACCGTACCGCCTCCCTTACACGACG 1080
 Qy 1367 CGGTTCTGATAGGCGCAGCGGCTGCTGGCGTGGTGCCCATGGGAATACCCCGCACCC 1426
 Db 1081 CGGTTCTGATAGGCGCAGCGGCTGCTGGCGTGGTGCCCATGGGAATACCCCGCACCC 1140
 Qy 1427 TCATGGGACACCCCGCTTCGAGGGTACACCTGCGCCAGGGGACGAGGCTTCCCGC 1486
 Db 1141 TCATGGGACACCCCGCTTCGAGGGTACACCTGCGCCAGGGGACGAGGCTTCCCGC 1200
 Qy 1487 TCCTGGCTCCATCCTGTCATGACCCCAACATCTTCAAGCACCCAGAAAGATTCAACCCAG 1546
 Db 1201 TCCTGGCTCCATCCTGTCATGACCCCAACATCTTCAAGCACCCAGAAAGATTCAACCCAG 1260
 Qy 1547 ACCGTTTCTGATGAGATGACGGTTTCAGGAAGCATGAGGCTTCTGCGCTTCTCT 1606
 Db 1261 ACCGTTTCTGATGAGATGACGGTTTCAGGAAGCATGAGGCTTCTGCGCTTCTCT 1320
 Qy 1607 TAGGGAAGCGTGTCTGCTTGGAGAGGCGCTGCAAAAGGAGGCTTCTCTCTTCTCA 1666
 Db 1321 TAGGGAAGCGTGTCTGCTTGGAGAGGCGCTGCAAAAGGAGGCTTCTCTCTTCTCA 1380
 Qy 1667 CCACCATCTTACAGCCTTCTCCTCGAGAGCCGTCGCCGCGACACCTGAGCTCA 1726
 Db 1381 CCACCATCTTACAGCCTTCTCCTCGAGAGCCGTCGCCGCGACACCTGAGGCTCA 1440
 Qy 1727 AGCCCAACCGTCAGTGGCTTTTCAACATTTCCCGACGCTTCCAGCTGCAAGTCCGTC 1786
 Db 1441 AGCCCAACCGTCAGTGGCTTTTCAACATTTCCCGACGCTTCCAGCTGCAAGTCCGTC 1500
 Qy 1787 CTGACCTTCACTCCACACGACACGATGAGGAGGCACTTGGAGTGGTGGTGC 1846
 Db 1501 CTGACCTTCACTCCACACGACACGATGAGGAGGCACTTGGAGTGGTGGTGC 1560
 Qy 1847 CCAGGACGGTGCCTCCAGCCTCAACAGTGGGATGACAGGGTTAATGTCTCCAGAGTGT 1906
 Db 1561 CCGGACGGTGCCTCCAGCCTCAACAGTGGGATGACAGGGTTAATGTCTCCAGAGTGT 1620
 Qy 1907 ACATCGAGCGACCAATTTACAGCCTGCAAGTGTGTTTCCGAGTCTGTCCCAAGGCG 1966
 Db 1621 ACATCGAGCGACCAATTTACAGCCTGCAAGTGTGTTTCCGAGTCTGTCCCAAGGCG 1680
 Qy 1967 CACACGCTCACTTGAATCATGCTGCTAAGATGCACACCGCACACCCATACCAACTACA 2026
 Db 1681 CACACGCTCACTTGAATCATGCTGCTAAGATGCACACCGCACACCCATACCAACTACA 1740

Qy 2027 AGGGCCACAAAGCAACTGCTGGGTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAA 2086
 Db 1741 AGGGCCACAAAGCAACTGCTGGGTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAA 1800
 Qy 2087 TCACAGACCATAGCAGGTAAACCAACCACTCCCTGTGATCTGACGCCCAACGCTGGGA 2146
 Db 1801 TCACAGACCATAGCAGGTAAACCAACCACTCCCTGTGATCTGACGCCCAACGCTGGGA 1860
 Qy 2147 GTCTGGTGTACCTTCAAGCCACAGAAACGCGCACACATGTTTCACAGCTCACACGCC 2206
 Db 1861 GTCTGGTGTACCTTCAAGCCACAGAAACGCGCACACATGTTTCACAGCTCACACGCC 1920
 Qy 2207 CTCTCATTTTCATCGAACTTCTCAGTGTCCCTGTCCCTGGTGGCTGGCACAGGAAACAGCA 2266
 Db 1921 CTCTCATTTTCATCGAACTTCTCAGTGTCCCTGTCCCTGGTGGCACAGGAAACAGCA 1980
 Qy 2267 TGCCCCCTCCGGGTCATGCCACCGAGACTGTGCTGTCTATGGCCCCCAACTCATGCT 2326
 Db 1981 TGCCCCCTCCGGGTCATGCCACCGAGACTGTGCTGTCTATGGCCCCCAACTCATGCT 2040
 Qy 2327 CCCTCTCTGGGTACACCACTCTCCAGCCTGTGACCCAGGATGTCACACACCCCAAC 2386
 Db 2041 CCCTCTCTGGGTACACCACTCTCCAGCCTGTGACCCAGGATGTCACACACCCCAAC 2100
 Qy 2387 CACTTGTCCACACAGCTACCCACGTACGACATCGTCTCGCTCCCGAGAGTATCTTCCCA 2446
 Db 2101 CACTTGTCCACACAGCTACCCACGTACGACATCGTCTCGCTCCCGAGAGTATCTTCCCA 2160
 Qy 2447 CTGAGACAGCGCGGCCCAACAGAGCAAGTCCCGACGACCTCTGCAACCTGAGGCGCTC 2506
 Db 2161 CTGAGACAGCGCGGCCCAACAGAGCAAGTCCCGACGACCTCTGCAACCTGAGGCGCTC 2220
 Qy 2507 AGTCACCCCTTTTAAAGCACCTGATTCTTACCAATGCAAAACACATCTCGGCTCTGCGATT 2566
 Db 2221 AGTCACCCCTTTTAAAGCACCTGATTCTTACCAATGCAAAACACATCTCGGCTCTGCGATT 2280
 Qy 2567 ATGCAACAGAGCTTTGGACATACGAGGACCTCTCAGACCGGAGGAACACCTGCCCCAACCC 2626
 Db 2281 ATGCAACAGAGCTTTGGACATACGAGGACCTCTCAGACCGGAGGAACACCTGCCCCAACCC 2340
 Qy 2627 AACAGTGTATTATGTAACACAGTGGAAAGCGGCGCTGTCGCCCTCCACACACATAC 2686
 Db 2341 AACAGTGTATTATGTAACACAGTGGAAAGCGGCGCTGTCGCCCTCCACACACATAC 2400
 Qy 2687 ACACCTCACTGATCTACAGCCCTGTTTCGGCGTCAGAGTCCCACTAGACCCAGTGAAGG 2746
 Db 2401 ACACCTCACTGATCTACAGCCCTGTTTCGGCGTCAGAGTCCCACTAGACCCAGTGAAGG 2460
 Qy 2747 GGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCACCCCTGTCAAGGAGTGAGCCGATCT 2806
 Db 2461 GGTTAGAGACCAAGTAGGGGCCAGTTTCCAATTCACCCCTGTCAAGGAGTGAGCCGATCT 2520
 Qy 2807 GAGCTTCTCTGACTTAAGGCTCGGCTTGGGAATTAAGTTGTTTCTGGCCTTTAGC 2866
 Db 2521 GAGCTTCTCTGACTTAAGGCTCGGCTTGGGAATTAAGTTGTTTCTGGCCTTTAGC 2580
 Qy 2867 CTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2905
 Db 2581 CTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2619

RESULT 4
 AX327331
 LOCUS
 DEFINITION Sequence 16 from Patent WO0179468.
 ACCESSION AX327331
 VERSION AX327331.1
 KEYWORDS GT:18097877
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AX327331 PAT 07-JAN-2002

Sequence 16 from Patent WO0179468.

GT:18097877

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

QY	1285	GTACGTGAGGAGCTGAATCGGGAGCTGGGGCTGGCCAGGACCAAGCTTAGGGGACCGT	1344
DB	1021	GTACGTGAGGAGCTGAATCGGGAGCTGGGGCTGGCCAGGACCAAGCTTAGGGGACCGT	1080
QY	1345	ACCGCCTCTCCTTACACCGAGCGCGTCTTGATCAGGCGCAGCGCTGCTGGCGCTGGTG	1404
DB	1081	ACCGCCTCTCCTTACACCGAGCGCGTCTTGATCAGGCGCAGCGCTGCTGGCGCTGGTG	1140
QY	1405	CCCATGGGAATACCCCGACCCCTCATCGGACCAACCCGCTTCCAGAGGTACACCTCGCC	1464
DB	1141	CCCATGGGAATACCCCGACCCCTCATCGGACCAACCCGCTTCCAGAGGTACACCTCGCC	1200
QY	1465	CAGGACGAGAGTCTTCCCGCTCTTGGCTCCTCATCTGATGACCCCAACATCTTCAAG	1524
DB	1201	CAGGACGAGAGTCTTCCCGCTCTTGGCTCCTCATCTGATGACCCCAACATCTTCAAG	1260
QY	1525	CACCCAGAAGATTCAACCCAGACCGTTTCTGTGATGCAGATGACGCTTTCAGGAAGCAT	1584
DB	1261	CACCCAGAAGTTCACCCAGACCGTTTCTGTGATGCAGATGACGCTTTCAGGAAGCAT	1320
QY	1585	GAGCGTTCCTGGCCTTCTCCTTAGGAAAGTGTCGCTTGGAGAGGCCTGGCAAAA	1644
DB	1321	GAGCGTTCCTGGCCTTCTCCTTAGGAAAGTGTCGCTTGGAGAGGCCTGGCAAAA	1380
QY	1645	GCGGAGTCTTCTCTTCTTTCACACCATCGTACAAGCCTTCTCCCTGGAGAGCCCGTGC	1704
DB	1381	GCGGAGTCTTCTCTTCTTTCACACCATCGTACAAGCCTTCTCCCTGGAGAGCCCGTGC	1440
QY	1705	CCGCGGACACCTGAGCCCTCAAGCCACACGTCAGTGGCCTTTTCAACATTCCTCCAGCC	1764
DB	1441	CCGCGGACACCTGAGCCCTCAAGCCACACGTCAGTGGCCTTTTCAACATTCCTCCAGCC	1500
QY	1765	TTCCAGTCTCAAGTTCGCTCCACTGACCTTCACTCCACCGCAGACCAAGTGAAGGAG	1824
DB	1501	TTCCAGTCTCAAGTTCGCTCCACTGACCTTCACTCCACCGCAGACCAAGTGAAGGAG	1560
QY	1825	GCAACTTGAAGTGGTGGTGCCAGAGCGGTGCTTCAGCCTCAACAGTGGGATGGAC	1884
DB	1561	GCAACTTGAAGTGGTGGTGCCAGAGCGGTGCTTCAGCCTCAACAGTGGGATGGAC	1620
QY	1885	AGGTTAACTCTCCAGAGTGTAACCTGACGAGGAGCCACATTTACAGCCCTGCAGTTGTT	1944
DB	1621	AGGTTAACTCTCCAGAGTGTAACCTGACGAGGAGCCACATTTACAGCCCTGCAGTTGTT	1680
QY	1945	TTCCGGAGTCTGCCACGGCCACACGCTCACTTGACTCATGCTGCTAGATGCAAC	2004
DB	1681	TTCCGGAGTCTGCCACGGCCACACGCTCACTTGACTCATGCTGCTAGATGCAAC	1740
QY	2005	CGCACACCCATACAACTACAGGGGCCAAAGCAAATGCTCTGGTTAGCTTTCCACAGA	2064
DB	1741	CGCACACCCATACAACTACAGGGGCCAAAGCAAATGCTCTGGTTAGCTTTCCACAGA	1800
QY	2065	CATAAATATAGTCCATCTGCAATCAAGACACATGCCAGGTAAACCAACTCCCTG	2124
DB	1801	CATAAATATAGTCCATCTGCAATCAAGACACATGCCAGGTAAACCAACTCCCTG	1860
QY	2125	GATCTCAGCCCCACAGTGGGAGTCTGGCTGTCACTTCAAGACCAAGAAAGGCCAC	2184
DB	1861	GATCTCAGCCCCACAGTGGGAGTCTGGCTGTCACTTCAAGACCAAGAAAGGCCAC	1920
QY	2185	ACATGTTACAGTCTACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCTG	2244
DB	1921	ACATGTTACAGTCTACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCTG	1980
QY	2245	GTCCCTGGCACAGGACAGCATGCCCTTCCGGGTGTCGCCACCGAGAGCTGCTGCT	2304
DB	1981	GTCCCTGGCACAGGACAGCATGCCCTTCCGGGTGTCGCCACCGAGAGCTGCTGCT	2040
QY	2305	GTCTATGGCCCCAACTCATGCTCCCTCTCTTTGGCTACACCACTCTCCAGCCTGTGACCA	2364
DB	2041	GTCTATGGCCCCAACTCATGCTCCCTCTCTTTGGCTACACCACTCTCCAGCCTGTGACCA	2100

QY	2365	CCGATGTCACACACACCCCGAACCTTTGTGTCACACAGCTACCCACGCTACGACATCGTCCT	2424
Db	2101	CCGATGTCACACACACCCCGAACCTTTGTGTCACACAGCTACCCACGCTACGACATCGTCCT	2160
QY	2425	GGCTCCCGCAGAGTATCTTCCACTGAGACACGGCGCCCCACACAGAGCAGTCCCGCAGC	2484
Db	2161	GGCTCCCGCAGAGTATCTTCCACTGAGACACGGCGCCCCACACAGAGCAGTCCCGCAGC	2220
QY	2485	CACCTCTGCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTCTACCAATGC	2544
Db	2221	CACCTCTGCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTCTACCAATGC	2280
QY	2545	AAACACATCTGGGTCTCGGATTTATGCACAGAGACTTTTGACATACAGGACCTCAGACC	2604
Db	2281	AAACACATCTGGGTCTCGGATTTATGCACAGAGACTTTTGACATACAGGACCTCAGACC	2340
QY	2605	GGAGGAACACTGCCCCAACCCCAACAGTGTATGTAAACAGTGTGAAGCGGCCCTG	2664
Db	2341	GGAGGAACACTGCCCCAACCCCAACAGTGTATGTAAACAGTGTGAAGCGGCCCTG	2400
QY	2665	CTGCCCCCTCCACACACACATACACTCACTGATCTACAGCCCCCTGTTGGCGCTCAGAGT	2724
Db	2401	CTGCCCCCTCCACACACACATACACTCACTGATCTACAGCCCCCTGTTGGCGCTCAGAGT	2460
QY	2725	CCCCACTAGACCCAGTGGAAAGGGTTAGAGACCAAGTAGGGCCAGCTTTCCAATTCACCC	2784
Db	2461	CCCCACTAGACCCAGTGGAAAGGGTTAGAGACCAAGTAGGGCCAGCTTTCCAATTCACCC	2520
QY	2785	TGTCAGGAGTAGCGCGGATCTACGTTCTCTGTGACTTAAAGGTCGGCTTGGGAATTA	2844
Db	2521	TGTCAGGAGTAGCGCGGATCTACGTTCTCTGTGACTTAAAGGTCGGCTTGGGAATTA	2580
QY	2845	AAGTTTGTCTTCTGGCCCTTTAGCCT	2868
Db	2581	AAGTTTGTCTTCTGGCCCTTTAGCCT	2604
RESULT 6			
AX552230			
LOCUS	AX552230	2604 bp	DNA
DEFINITION	Sequence 139 from Patent WO0162927.		linear
ACCESSION	AX552230		
VERSION	AX552230.1	GI:25896468	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Banville,S.C., Greensawalt,L.B., Lincoln,S.E., Stockdreher,T.K., Amshew,S., Chang,S.C., Chen,W., D'Sa,S.A., Dam,T.C., Liu,T.F., Rosen,B.H., Russo,F.D., Spiro,P.A., Bradley,D.L., Chen,A., Cohen,H.J., Daffo,A., Daniels,S.E., Dufour,G.E., Flores,V., Fong,W.T., Hodgson,D.M., Jackson,S., Jones,A.L., Panzer,S., Roseberry,A.M., Shah,P., Wright,R.J., Yap,P.E., Yu,J.Y., Braticher,S.R., Chalup,M.S., Dahl,C.R. and Hillman,J.L.		
TITLE	Polypeptides and corresponding polynucleotides for diagnostics and therapeutics		
JOURNAL	Patent: WO 0162927-A 139 30-AUG-2001;		
FEATURES	Incyte Genomics, Inc. (US)		
source	Location/Qualifiers		
	1..2604		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN	/note="Incyte ID No: LG:337818.2:20000FEB18"		
Query Match	88.2%;	Score 2596;	DB 6; Length 2604;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 2599;	Conservative 0;	Mismatches 5;	Indels 0; Gaps 0;
QY	265	CGGAGCGCTGGGAGGAGGAGGAGCGGACCTCGCAGATGGAGGCGACCGGACCTGG	324


```
QY 2485 CACCTCTGCACTGAGCCCTCAGTACACCCCTTTTAAAGACCCCTGATTTCTACCAATGC 2544
Db 2221 CACCTCTGCACTGAGCCCTCAGTACACCCCTTTTAAAGACCCCTGATTTCTACCAATGC 2280
QY 2545 AAACACATCTGGTCTGCGATTATGACAGAGACTTTGGACATACAGAGACCCCTCAGACC 2604
Db 2281 AAACACATCTGGTCTGCGATTATGACAGAGACTTTGGACATACAGAGACCCCTCAGACC 2340
QY 2605 GGAGGAACCTGCGCAACCCCAACACAGTGTATGTAACACAGTGGAAAGCGCCCTG 2664
Db 2341 GGAGGAACCTGCGCAACCCCAACACAGTGTATGTAACACAGTGGAAAGCGCCCTG 2400
QY 2665 CTGCCCCCTCCACACACATACACACTCAGTATGTAACACAGTGGAAAGCGCCCTG 2724
Db 2401 CTGCCCCCTCCACACACATACACACTCAGTATGTAACACAGTGGAAAGCGCCCTG 2460
QY 2725 CCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGCCAGTTTCCAAATTCACCC 2784
Db 2461 CCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGCCAGTTTCCAAATTCACCC 2520
QY 2785 TGTGAGGAGTACCGGAGTCTGAGCTTCTGTGACATTAAGGTTCGGCTTGGGAATTA 2844
Db 2521 TGTGAGGAGTACCGGAGTCTGAGCTTCTGTGACATTAAGGTTCGGCTTGGGAATTA 2580
QY 2845 AAGTTTCTTCTGGCCCTTAGCCT 2868
Db 2581 AAGTTTCTTCTGGCCCTTAGCCT 2604

RESULT 7
AX876954
LOCUS 2513 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 11859 from Patent EP1074617.
ACCESSION AX876954
VERSION AX876954.1 GI:40031690
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primers for synthesising full-length cDNA and their use
Patent: EP 1074617-A 11859 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1..2513
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
56..1750
/notes="unnamed protein product"
/codon_start=1
/protein_id="CA889506.1"
/db_xref="GI:40031691"
/translation="MEATGTLWALLALALLLLTLSTGRAGHLPPGPTPLPLGN
LIQLRGALYGLMRSLKSKYGVFTIYLGFWPVLVQGEAVREALGQAEFSGRG
TVAMLEGFDGHGVFFSNGERWQLRFTMLALDLGMKREGEELIQAEACLVTFF
QTEGPPFPDPSLLLAQATSNVCSLLFLRFSVEDKEFOAVVRAAGTLLGVSSQSGQ
TYEMFQWFLRPLPGPHQKLLHVVSTIAFTVQVQHQGNLDASGPARDLVAPFLKM
AOEONPGTEFTKNMLTVIYLLFAGTIVSTVGYTLMLLMKYTHVQKRWREELNR
ELGAGAPSLGDRTRLPYDVALHEAQRLLALVPMGIPRTLMRTTFRGYTLPPQSTEV
FPLGSLIHDNFIKHEPEFNDFLDADGRFKHEAFLPFLSLGRVCLGEGELAKEL
FIFFTIIQAFSLSPEDTSLSKPTVSGLENIPTASTVGMDRVNVSVYTAGSHY
TPAVFRSLSHGPHALHTAAKMHNRTPIHNVKHKATAGLAFHRHKYSPSALTST"

Query Match 81.5%; Score 2399.4; DB 6; Length 2513;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 2512; Conservative 0; Mismatches 1; Indels 102; Gaps 1;

ORIGIN
```

Db	1081	AAGCCTAGGGGACCGTATCCCGCCTTCCCTTACACCGACGGGTTCTGCATAGGCGCAGCG	1140
Qy	1389	GCTGCTGGCGTGGTGCCCATGGGAATACCCCGCACCTCATCGGAGCACACCGGCTTCG	1448
Db	1141	GCTGCTGGCGTGGTGCCCATGGGAATACCCCGCACCTCATCGGAGCACACCGGCTTCG	1200
Qy	1449	AGGGTACACCTGCGCCAGGCGACGAGGTCTTCCCCTCTCTTGGGTCTCATCTGTCATGA	1508
Db	1201	AGGGTACACCTGCGCCAGGCGACGAGGTCTTCCCCTCTCTTGGGTCTCATCTGTCATGA	1260
Qy	1509	CCCCAACATCTTCAAGCACCCAGGAAGTTCAAACCCAGACCGTTTCTTGGATGCAATGG	1568
Db	1261	CCCCAACATCTTCAAGCACCCAGGAAGTTCAAACCCAGACCGTTTCTTGGATGCAATGG	1320
Qy	1569	ACGGTTGAGGAAGCATGAGGCGTTCCTGCCCTCTCTTAGGGAAGCGGTCTGCGCTTGG	1628
Db	1321	ACGGTTGAGGAAGCATGAGGCGTTCCTGCCCTCTCTTAGGGAAGCGGTCTGCGCTTGG	1380
Qy	1629	AGGGGCTGTGCAAAAGCGGAGCTTCTTCTCTTCCACCAACCATCTCAAGGCCCTTCTC	1688
Db	1381	AGAGGGCTGTGCAAAAGCGGAGCTTCTTCTCTTCCACCAACCATCTCAAGGCCCTTCTC	1440
Qy	1689	CCTGGAGAGCCGTGCGCGCGGACACCTTGAGCCTCAAGCCCAACGTCAGTGGCCTTTT	1748
Db	1441	CCTGGAGAGCCGTGCGCGCGGACACCTTGAGCCTCAAGCCCAACGTCAGTGGCCTTTT	1500
Qy	1749	CAACATTTCCCGACGCTTCCAGCTGCGAAGTCCGTCCCACTGACCTTCACTCCACCAGCA	1808
Db	1501	CAACATTTCCC-----	1510
Qy	1809	GACCAGATGAAGGAAGGCAACTTGGAAATGGTGGTGCCAGGACGGTGCCTCCAGCCTC	1868
Db	1511	-----CGAGCCTC	1518
Qy	1869	AACAGTGGGCATGGACAGGGTTAATGTCTCAGAGTGTACACTGCAGGCAGCCACATTTA	1928
Db	1519	AACAGTGGGCATGGACAGGGTTAATGTCTCAGAGTGTACACTGCAGGCAGCCACATTTA	1578
Qy	1929	CACGGCTGCAGTTGTTTTCGGAGTCTGTCCACGGCCACACGCTCACTTGACTCATGC	1988
Db	1579	CACGGCTGCAGTTGTTTTCGGAGTCTGTCCACGGCCACACGCTCACTTGACTCATGC	1638
Qy	1989	TGCTAAGATGCACAACCGCACACCCATACAACTACAAGGGCCACAAAGCACTGCTGG	2048
Db	1639	TGCTAAGATGCACAACCGCACACCCATACAACTACAAGGGCCACAAAGCACTGCTGG	1698
Qy	2049	GTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCACAGCACATAGCCAGTAA	2108
Db	1699	GTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCACAGCACATAGCCAGTAA	1758
Qy	2109	CCACACAACTCCCTGGATCTGACGCCACACGTGGGAGTCTGGCTGCACCTTCAAG	2168
Db	1759	CCACACAACTCCCTGGATCTGACGCCACACGTGGGAGTCTGGCTGCACCTTCAAG	1818
Qy	2169	CCACAGAAACGGCCACACATGTTTCAAGCTCACACGCCCTCTCCATTTATCGAACTTCTC	2228
Db	1819	CCACAGAAACGGCCACACATGTTTCAAGCTCACACGCCCTCTCCATTTATCGAACTTCTC	1878
Qy	2229	AGTGTCCCTGTCCCTGGTGCCTGGGCACAGGGAACAGATGCCCTCCGGGTCATGCCA	2288
Db	1879	AGTGTCCCTGTCCCTGGTGCCTGGGCACAGGGAACAGATGCCCTCCGGGTCATGCCA	1938
Qy	2289	CCACAGACTGTGCGTGTCTATGGCCCAACTCATGTCTCCCTCTCTTGGCTACACCACTC	2348
Db	1939	CCACAGACTGTGCGTGTCTATGGCCCAACTCATGTCTCCCTCTCTTGGCTACACCACTC	1998
Qy	2349	TCCAGACCTGTGACCAACCGATGTCCACACACCCCCAACCACTGTCCACACAGCTACCCA	2408
Db	1999	TCCAGACCTGTGACCAACCGATGTCCACACACCCCCAACCACTGTTCACACAGCTACCCA	2058
Qy	2409	CGTAGACATGCTCTGGCTCCCGCAGATATCTTCCCACTGAGACACGCGGCCCCCAACAG	2468

Db	2059	CGTAGACATCGTCCCGCTCCCGACAGATATCTTCCCACTGAGACACGCGCGCCCAACAG	2111
Qy	2469	AGGCACAGTCCCGACGACCACTCTGCAACTGACGCGCTTCACTGACCCCTTTTAAAGCAACC	2528
Db	2119	AGGCACAGTCCCGACGACCACTCTGCAACTGACGCGCTTCACTGACCCCTTTTAAAGCAACC	2178
Qy	2529	TGATTTACCAATGCAAAACATCTCTGGTCTGGGATTTATGCACAGAGACTTTGGACATA	2588
Db	2179	TGATTTACCAAAATGCAAAACATCTCTGGTCTGGGATTTATGCACAGAGACTTTGGACATA	2238
Qy	2589	CGAGGACCTCAGACCGGAGGACACCTGCGCCACCCCAACACAGCTGCTTATGTAACCAACG	2648
Db	2239	CGAGGACCTCAGACCGGAGGACACCTGCGCCACCCCAACACAGCTGCTTATGTAACCAACG	2298
Qy	2649	TGAAAGCGGCGCTGCTGCTGCGCTTCCACACACACATACACACTCACTGATCTACAGCCCC	2708
Db	2299	TGAAAGCGGCGCTGCTGCTGCGCTTCCACACACACATACACACTCACTGATCTACAGCCCC	2358
Qy	2709	TGTTGGCGGTGAGAGTCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGGCC	2768
Db	2359	TGTTGGCGGTGAGAGTCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGGCC	2418
Qy	2769	AGTTTCCAATTACCCCTGTGAGGAGTGAAGCGGATCTGACGTTCTCTGACTTAAGGG	2828
Db	2419	AGTTTCCAATTACCCCTGTGAGGAGTGAAGCGGATCTGACGTTCTCTGACTTAAGGG	2478
Qy	2829	TCCGGCTTGGAAATAAAGTTTGTCTGGCGTTT	2863
Db	2479	TCCGGCTTGGAAATAAAGTTTGTCTGGCGTTT	2513
RESULT 8			
BD156406		2513 bp DNA linear	PAT 17-JAN-2003
LOCUS	BD156406	Primer for synthesizing full-length cDNA and use thereof.	
DEFINITION	BD156406.1	GI:27862164	
ACCESSION	BD156406.1	GI:27862164	
VERSION	JP 2002191363-A/11249		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.		
TITLE	Primer for synthesizing full-length cDNA and use thereof		
JOURNAL	Patent: JP 2002191363-A 11249 09-JUL-2002; HELIX RESEARCH INSTITUTE		
COMMENT	OS Homo sapiens (human) PN JP 2002191363-A/11249 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO, PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, PI KEIICHI NAGAI, TETSUJI OTSUKI PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/00, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC		
FEATURES	source	Primer for synthesizing full-length cDNA and use thereof FH Key	
	FT CDS	Location/Qualifiers	
		1..2513	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match	81.5%;	Score 2399.4;	DB 6; Length 2513;
Best Local Similarity	96.1%;	Freq. No. 0;	
Matches 2512; Conservative	0;	Mismatches	1; Indels 102; Gaps 1

QY 969 GTTCTCTGGTTCCTGCGGCCCTGCCAGGCCCCCACAAGCAGCTCCTCCACCAAGTCAG 1028
DB 721 GTTCTCTGGTTCCTGCGGCCCTGCCAGGCCCCCACAAGCAGCTCCTCCACCAAGTCAG 780
QY 1029 CACCTTGGCTGCTTCACAGTCCGGCAGGTGAGCAGCACCAGGGGAACTGGATGCTTC 1088
DB 781 CACCTTGGCTGCTTCACAGTCCGGCAGGTGAGCAGCACCAGGGGAACTGGATGCTTC 840
QY 1089 GGGCCCCCAGTGCACCTTGTGATGCTTCCTGCTGAAGATGGCAGAGGAGAAACAAA 1148
DB 841 GGGCCCCCAGTGCACCTTGTGATGCTTCCTGCTGAAGATGGCAGAGGAGAAACAAA 900
QY 1149 CCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTATTTGCTGTTTC 1208
DB 901 CCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTATTTGCTGTTTC 960
QY 1209 TGGGACGATGACGCTGACGACACAGGTGCGGTATACCCCTCTGCTCTGATGAATACCC 1268
DB 961 TGGGACGATGACGCTGACGACACAGGTGCGGTATACCCCTCTGCTCTGATGAATACCC 1020
QY 1269 TCATGTCCTCAAGTGGGTACCTGAGGAGCTGAATCGGAGCTGGGGCTGSCCAGGCACC 1328
DB 1021 TCATGTCCTCAAGTGGGTACCTGAGGAGCTGAATCGGAGCTGGGGCTGSCCAGGCACC 1080
QY 1329 AAGCCTAGGGGACCGGTACCCCGCTCCCTTACACGACGCGGTTCTGCAATGAGGCGCAGC 1388
DB 1081 AAGCCTAGGGGACCGGTACCCCGCTCCCTTACACGACGCGGTTCTGCAATGAGGCGCAGC 1140
QY 1389 GCTGCTGGCGTGTGCGCATGCGGAATACCCCGCACCTCATGCGGACCAACCGCTTCCG 1448
DB 1141 GCTGCTGGCGTGTGCGCATGCGGAATACCCCGCACCTCATGCGGACCAACCGCTTCCG 1200
QY 1449 AGGTGACACCTGCCCCAGGCGAGGCTTCCCTCCCTCTGCTGCTCCATCTCGATGA 1508
DB 1201 AGGTGACACCTGCCCCAGGCGAGGCTTCCCTCCCTCTGCTGCTCCATCTCGATGA 1260
QY 1509 CCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTGATGACAGATGG 1568
DB 1261 CCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTGATGACAGATGG 1320
QY 1569 ACGTTTCAAGAGATGAGGCTTCTGCTTCTTCAAGGAGGCTGCTGCTGCTGG 1628
DB 1321 ACGTTTCAAGAGATGAGGCTTCTGCTTCTTCAAGGAGGCTGCTGCTGCTGG 1380
QY 1629 AGAGGCTGCGCAAGGAGCTTCTTCTTCTTCAACACCATCTCAAGCCTTCTC 1688
DB 1381 AGAGGCTGCGCAAGGAGCTTCTTCTTCTTCAACACCATCTCAAGCCTTCTC 1440
QY 1689 CTGGAGAGCCGTCGCGCGGACACCTGAGCTCAAGCCCAACCTCAAGTGGCCTTTT 1748
DB 1441 CTGGAGAGCCGTCGCGCGGACACCTGAGCTCAAGCCCAACCTCAAGTGGCCTTTT 1500
QY 1749 CAACATCTCCCGAGCTTCCAGCTGCAAGTCCGTCCTCACTGACCTTCACTCCACCAACGCA 1808
DB 1501 CAACATCTCC----- 1510
QY 1809 GACCAGATGAAGGAAGCACTTGGAGTGGTGGTCCCGAGGAGTGGCTCCAGCCTC 1868
DB 1511 -----CCAGCCTC 1518
QY 1869 AACAGTGGGATGGACAGAGGTAAATGCTCCAGAGTGTACACTGCAGGAGCCACATTTA 1928
DB 1519 AACAGTGGGATGGACAGAGGTAAATGCTCCAGAGTGTACACTGCAGGAGCCACATTTA 1578
QY 1929 CAGCCTGCAAGTTGTTTCCGGAGTGTGCTCCAGGCCCCACAGCTCACTTGACTCATGC 1988
DB 1579 CAGCCTGCAAGTTGTTTCCGGAGTGTGCTCCAGGCCCCACAGCTCACTTGACTCATGC 1638
QY 1989 TGCTAGATGCACACCGCACCCATACACAACTACAAGGGCCACAAAGCACTGCTGG 2048
DB 1639 TGCTAGATGCACACCGCACCCATACACAACTACAAGGGCCACAAAGCACTGCTGG 1698
QY 2049 GTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGGTAA 2108

DB 1699 GTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGGTAA 1758
QY 2109 CCCACCAACTCCCTTGGATCTGCGAGCCCAACAGTGGGAGTCTGGCTGTACACCTTCCACAG 2168
DB 1759 CCCACCAACTCCCTTGGATCTGCGAGCCCAACAGTGGGAGTCTGGCTGTACACCTTCCACAG 1818
QY 2169 CCACAGAAACGGCCACACATGTTTCCAGCTCACAGCCCTCTCCATTTCCATCGAACTTCTC 2228
DB 1819 CCACAGAAACGGCCACACATGTTTCCAGCTCACAGCCCTCTCCATTTCCATCGAACTTCTC 1878
QY 2229 AGTGTCCCTGTCCTGCTGCTGCGACAGGAAACAGATGCCCTCCGGGGTATGCA 2288
DB 1879 AGTGTCCCTGTCCTGCTGCTGCGACAGGAAACAGATGCCCTCCGGGGTATGCA 1938
QY 2289 CCCAGAGACTGTGCTGTCTATGCGCCCAACTCATGCTCCCTCTCTTGGCTACACCACTC 2348
DB 1939 CCCAGAGACTGTGCTGTCTATGCGCCCAACTCATGCTCCCTCTCTTGGCTACACCACTC 1998
QY 2349 TCCAGAGCTGTGACACACCGATGTCCACACCCCAACCACTTGTCCACACAGCTTACCCA 2408
DB 1999 TCCAGAGCTGTGACACACCGATGTCCACACCCCAACCACTTGTCCACACAGCTTACCCA 2058
QY 2409 CGTAGACATGCTGCTGCTGCTGCGACAGATATCTTCCCACTGAGACACGCGCCGCCACAG 2468
DB 2059 CGTAGACATGCTGCTGCTGCTGCGACAGATATCTTCCCACTGAGACACGCGCCGCCACAG 2118
QY 2469 AGGCACAGTCTCCCAACCACTCTGCAACTGCGAGCCCTCAGTCAACCCCTTTTAAAGCACC 2528
DB 2119 AGGCACAGTCTCCCAACCACTCTGCAACTGCGAGCCCTCAGTCAACCCCTTTTAAAGCACC 2178
QY 2529 TGATTCTACCAATGCAACACATCTGGTCTGCGATTATGCAACAGAGACTTTGACATA 2588
DB 2179 TGATTCTACCAATGCAACACATCTGGTCTGCGATTATGCAACAGAGACTTTGACATA 2238
QY 2589 CGAGGACCTCAGACCGGAGGACACCTGCCCAACCCCAACCCCAACCTGTTTATGTAACACG 2648
DB 2239 CGAGGACCTCAGACCGGAGGACACCTGCCCAACCCCAACCCCAACCTGTTTATGTAACACG 2298
QY 2649 TGGAAAGCGGCCCTGCTGCGCCCTCCACACACATACACTCACTGATCTACAGCCCC 2708
DB 2299 TGGAAAGCGGCCCTGCTGCGCCCTCCACACACATACACTCACTGATCTACAGCCCC 2358
QY 2709 TGTTCGGCGTACAGTCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGGCC 2768
DB 2359 TGTTCGGCGTACAGTCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGGCC 2418
QY 2769 AGTTTCCAATTCACCTGCTCAGGAGTGAGCGGATCTGAGCTTCTTGTGACTTAAGG 2828
DB 2419 AGTTTCCAATTCACCTGCTCAGGAGTGAGCGGATCTGAGCTTCTTGTGACTTAAGG 2478
QY 2829 TCCGGCTTGGGAATTAAGTTTCTGCGCCCTT 2863
DB 2479 TCCGGCTTGGGAATTAAGTTTCTGCGCCCTT 2513

RESULT 10
AX780433

LOCUS
DEFINITION
AX780433
ACCESSION
VERSION
AX780433.1
KEYWORDS

Sequence 2590 from Patent WO03039443.
GI:32697427

2261 bp DNA linear PAT 14-JUL-2003

Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
AUTHORS
Haerlath, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
Dugas, M., Bils, R., Brors, B. and Mergenthaler, S.
TITLE
Novel genetic markers for leukemias
JOURNAL
Patent: WO 03039443-A 2590 15-MAY-2003;
Deutsches Krebsforschungszentrum (DKFZ);


```

|||||
1981 1GCCACCAGAGACTGTCGTCGTCATATGCGCCCACTCATGTCCTCTCTCTGCTACAC 2040
Db
QY 2344 CACTCTCCAGCTGTGACCAACCGATGTCACACACCCCAACCACTGTGTCACACAGCT 2403
Db
2041 CACTCTCCAGCTGTGACCAACCGATGTCACACACCCCAACCACTGTGTCACACAGCT 2100
QY 2404 ACCACGTCACACATCGTCCTGGCTCCCGACAGTATCTCCACACGAGACGCGCGCC 2463
Db
2101 ACCACGTCACACATCGTCCTGGCTCCCGACAGTATCTCCACGAGACGCGCGCC 2160
QY 2464 CACAGAGCACAGTCCCGACGACCTCTGCACTGACGCGCTCAGTCACCCCTTTTAAAG 2523
Db
2161 CACAGAGCACAGTCCCGACGACCTCTGCACTGACGCGCTCAGTCACCCCTTTTAAAG 2220
QY 2524 CACCTGA 2531
Db
2221 CACCATGA 2228

RESULT 11
AF335278 2261 bp mRNA linear PRI 28-FEB-2001
LOCUS Homo sapiens cytochrome P450 2S1 (CYP2S1) mRNA, complete cds.
ACCESSION AF335278
VERSION AF335278.1 GI:13161183
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2261)
Rylander,T., Neve,E.P., Ingelman-Sundberg,M. and Oscarson,M.
Identification and tissue distribution of the novel human
cytochrome P450 2S1 (CYP2S1)
Biochem. Biophys. Res. Commun. 281 (2), 529-535 (2001)
MEDLINE 21092856
PUBMED 1181079
REFERENCE 2 (bases 1 to 2261)
Rylander,T., Neve,E., Ingelman-Sundberg,M. and Oscarson,M.
Direct Submission
Submitted (10-JAN-2001) Institute of Environmental Medicine,
Molecular Toxicology, Karolinska Institute, Nobelsv. 13, Stockholm
17177, Sweden
FEATURES
source Location/Qualifiers
1..2261
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
1..2261
/gene="CYP2S1"
1..1515
/gene="CYP2S1"
/codon_start=1
/product="cytochrome P450 2S1"
/protein_id="AAK13498.1"
/db_xref="GI:13161184"
/translation="MEATGWALLLALLLLTLASGTRARGHLPPGPTPLLLGN
LLQLRGLALYGLMRSLKYGVPVFTIYLGPEVVLVGOEAVREALGGQAEFFSGR
TVAMLEGTDFGHVFFSNGERWRLKFTMLALDLMGKRGEEELIQAEARCLVET
QYGEFPFDPSSLQAQSNVCSLLFLRFSVEDKEFOAVVRAAGTLLGVSSQGG
TYMFSWFLRPLPGPKQLLHVSTLAAFTVRVQQHQNLDASGARDLIVDFLNR
AQEQNPGEFTKNMLTVIYLLFAGTMVSTVGTLLLLMKYPHVQKWRBELNR
ELAGQAPSLGDRILRUPYTDVLUHEAQRLLALPMGIPRTLMRTFRGYTLPGQTEV
FPLGLSLHDPNIKFPEEFENPDRFLDAGFRKHEAFPLFSIGKRVCLGEGIAKEL
FLFETTLQAFSLSEPCPDLSLKPTVSLFNIPAFQLQVPTDLHSTQIR"

Query Match 75.8%; Score 2223.2; DB 9; Length 2261;
Best Local Similarity 99.94; Pred. No. 0;
Matches 2225; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
ORIGIN

```

```

304 ATGAGGGGACACGGCACCTTGCGCGCTGCTGTGGCGCTGGCGCTGCTCTCTGTGCTGACG 363
1 ATGAGGGGACACGGCACCTTGCGCGCTGCTGTGGCGCTGGCGCTGCTCTCTGTGCTGACG 60
QY 364 CTGGCGCTGTCTGGGACACAGGCGCCGAGGCGCACCTGCGCCCGCGGCGCCACGCGGTACCA 423
Db
61 CTGGCGCTGTCTGGGACACAGGCGCCGAGGCGCACCTGCGCCCGCGGCGCCACGCGGTACCA 120
QY 424 CTGTGGGAAAACCTCTCTGAGCTACGGCCCGGGCGCTGTATTTCAGGGCTCATGCGGCTG 483
Db
121 CTGTGGGAAAACCTCTCTGAGCTACGGCCCGGGCGCTGTATTTCAGGGCTCATGCGGCTG 180
QY 484 AGTAAGAAGTAGGACCGGTGTTCACATCTACCTGGAACCCCTGCGCGCTGTGTGTGTGTC 543
Db
181 AGTAAGAAGTAGGACCGGTGTTCACATCTACCTGGAACCCCTGCGCGCTGTGTGTGTGTC 240
QY 544 CTGTTGGGACAGGAGCTGCGGGAGGCGCCCTGGGAGGTGAGGCTGAGGAGTTACGCGC 603
Db
241 CTGTTGGGACAGGAGCTGCGGGAGGCGCCCTGGGAGGTGAGGAGTTACGCGC 300
QY 604 CGGGAAACCTGAGCATGCTGGAAGGAGCTTTTATGAGCCATGGGGTTTCTCTCTCCAAAC 663
Db
301 CGGGAAACCTGAGCATGCTGGAAGGAGCTTTTATGAGCCATGGGGTTTCTCTCTCCAAAC 360
QY 664 GGGGACGCGTGGAGGAGCTGAGGAAGTTTACATGCTTGTCTGTGCGGGAACCTGGGCGATG 723
Db
361 GGGGACGCGTGGAGGAGCTGAGGAAGTTTACATGCTTGTCTGTGCGGGAACCTGGGCGATG 420
QY 724 GGGGACGAGAAAGCGGAGGAGCTGATCCAGCGGAGGCGCCGCTGTGTGCGGAGGAGCATTC 783
Db
421 GGGGACGAGAAAGCGGAGGAGCTGATCCAGCGGAGGCGCCGCTGTGTGCGGAGGAGCATTC 480
QY 784 CAGGGACAGAAAGAACCCCATTCGATTCCTCTGCTGCGGCCAGGCCACCTCCAAAC 843
Db
481 CAGGGACAGAAAGAACCCCATTCGATTCCTCTGCTGCGGCCAGGCCACCTCCAAAC 540
QY 844 GTAGTCTGCTCCCTCTCTTTGGCTCGGCTTCTCTATGAGGATAAGGAGTTCCAGGCC 903
Db
541 GTAGTCTGCTCCCTCTCTTTGGCTCGGCTTCTCTATGAGGATAAGGAGTTCCAGGCC 600
QY 904 GTGTCCTGGGACGCTGCTGCTGAGTACCTCTGCGGAGTACCTCCAGGGGGGTGACACCTAC 963
Db
601 GTGTCCTGGGACGCTGCTGCTGAGTACCTCTGCGGAGTACCTCCAGGGGGGTGACACCTAC 660
QY 964 GAGATGTTCTCTGGTTCTCTGCGGCCCTCTGCGGCCCGCCACAGCAGCTCTCTCCACAC 1023
Db
661 GAGATGTTCTCTGGTTCTCTGCGGCCCTCTGCGGCCCGCCACAGCAGCTCTCTCCACAC 720
QY 1024 GTGACGACCTTGGCTGCTTACAGTCCGCGAGGTGAGCAGCAGCAGGCGGAACTGGAT 1083
Db
721 GTGACGACCTTGGCTGCTTACAGTCCGCGAGGTGAGCAGCAGCAGGCGGAACTGGAT 780
QY 1084 GCTTTCGGGCGCCCGCAGTACCTTGTGATGCTTCTCTGCTGAAGATGGGCACAGGAGAA 1143
Db
781 GCTTTCGGGCGCCCGCAGTACCTTGTGATGCTTCTCTGCTGAAGATGGGCACAGGAGAA 840
QY 1144 CAAAAACCGGACAGAAATTCACCAACAGAAATGCTGATGACAGTCAATTTATTTCTG 1203
Db
841 CAAAAACCGGACAGAAATTCACCAACAGAAATGCTGATGACAGTCAATTTATTTCTG 900
QY 1204 TTTGCTGGGACGATGAGGTGAGCAGCAGCGTCGGCTATACCTCTCTGCTCTCTGATGAAA 1263
Db
901 TTTGCTGGGACGATGAGGTGAGCAGCAGCGTCGGCTATACCTCTCTGCTCTCTGATGAAA 960
QY 1264 TACCTCTCATGTCCAAAGTGGGTACGTGAGGAGCTCAATCGGGAGCTGGGGGTGCCAG 1323
Db
961 TACCTCTCATGTCCAAAGTGGGTACGTGAGGAGCTCAATCGGGAGCTGGGGGTGCCAG 1020
QY 1324 GCACCAAGCCTAGGGACCGTATACCGCCCTCCCTTACACCGACCGCGTCTGCAATGAGCG 1383
Db
1021 GCACCAAGCCTAGGGACCGTATACCGCCCTCCCTTACACCGACCGCGTCTGCAATGAGCG 1080
QY 1384 CAGCGGCTGCTGGCGCTGGTGCCCTAGGGAATACCCCGCACCCCTCATGCGGACACCCCG 1443

```

[illegible]

Job time : 11160.1 secs

Db	71554	GAAGTGTGGGTGCGCGGAGCGTCCCTCCAGCTCAACAGTGGGCATGGACAGGGTTA	71495
Qy	1892	ATGCTCCAGAGTGACACTGACGAGCAGCCATTTACAGCCTGAGTGTTCGGGA	1951
Db	71494	ATGCTCCAGAGTGACACTGACGAGCAGCCATTTACAGCCTGAGTGTTCGGGA	71435
Qy	1952	GTCTGTCCACGGCCCAACAGCTCACTTGACTCATGCTGCTAAGATGACACACCGCAC	2011
Db	71434	GTCTGTCCACGGCCCAACAGCTCACTTGACTCATGCTGCTAAGATGACACACCGCAC	71375
Qy	2012	CCATACACAACCTAACAGGGCCCAAAAGCAACTGTCTGGGTAGCTTTCCACAGACATAAT	2071
Db	71374	CCATACACAACCTAACAGGGCCCAAAAGCAACTGTCTGGGTAGCTTTCCACAGACATAAT	71315
Qy	2072	ATAGTCCATCTGCAATCACAGACATAGCCAGGTAAAGCCCAACTCCCTGGATCTGC	2131
Db	71314	ATAGTCCATCTGCAATCACAGACATAGCCAGGTAAAGCCCAACTCCCTGGATCTGC	71255
Qy	2132	AGCCCAACAGTGGAGTCTGGCTGTCACTTCAAGCCACAGAAAGCCACACATGTT	2191
Db	71254	AGCCCAACAGTGGAGTCTGGCTGTCACTTCAAGCCACAGAAAGCCACACATGTT	71195
Qy	2192	CACAGCTCACAGGCTCTCAATTCATGCAACTTCTCAGTGTCCCTGGTGCCTG	2251
Db	71194	CACAGCTCACAGGCTCTCAATTCATGCAACTTCTCAGTGTCCCTGGTGCCTG	71135
Qy	2252	GCACAGGGAACAGCATGCCCTCCGGGTGATGCCACCCAGAGACTGTGCTGTCTATG	2311
Db	71134	GCACAGGGAACAGCATGCCCTCCGGGTGATGCCACCCAGAGACTGTGCTGTCTATG	71075
Qy	2312	GCCTCAACTCATGCTCCCTCTCTTGGGTACACCACTCTCCAGGCTGTGACACCGATGT	2371
Db	71074	GCCTCAACTCATGCTCCCTCTCTTGGGTACACCACTCTCCAGGCTGTGACACCGATGT	71015
Qy	2372	CCACACACCCCAACCACTTGTCCACACAGTACCCACGTACGATCGTCCCTGCC	2431
Db	71014	CCACACACCCCAACCACTTGTCCACACAGTACCCACGTACGATCGTCCCTGCC	70955
Qy	2432	CAGAGTATCTCCCACTGACACAGCGCCGCCACAGAGGACAGTCCCGAGCCCTCT	2491
Db	70954	CAGAGTATCTCCCACTGACACAGCGCCGCCACAGAGGACAGTCCCGAGCCCTCT	70895
Qy	2492	GCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTCTACCAATGCAACACA	2551
Db	70894	GCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTCTACCAATGCAACACA	70835
Qy	2552	TCTGGGTCTGCGATTATGCAAGAGACTTTGGACATACGAGGACCCCTCAGACCGGAGAA	2611
Db	70834	TCTGGGTCTGCGATTATGCAAGAGACTTTGGACATACGAGGACCCCTCAGACCGGAGAA	70775
Qy	2612	CACCTGCCCAACCCCAACAGCTTATGTAACCACTGGAAGAGCGCCCTCTGTCGCC	2671
Db	70774	CACCTGCCCAACCCCAACAGCTTATGTAACCACTGGAAGAGCGCCCTCTGTCGCC	70715
Qy	2672	TCACACACATACACACTCACTGATCTACAGCCCTGTTCCGGCGTCAGAGTCCCACT	2731
Db	70714	TCACACACATACACACTCACTGATCTACAGCCCTGTTCCGGCGTCAGAGTCCCACT	70655
Qy	2732	AGACCCAGTGGAGGGGTAGAGACCAAGTAGGGGCAAGTTCCCAATTCACCTGTGAG	2791
Db	70654	AGACCCAGTGGAGGGGTAGAGACCAAGTAGGGGCAAGTTCCCAATTCACCTGTGAG	70595
Qy	2792	GAGTGAGCGGATCTCAGCTTCCCTGCTGACTTAAGGCTCGGCTTGGGAATTAAGTTTG	2851
Db	70594	GAGTGAGCGGATCTCAGCTTCCCTGCTGACTTAAGGCTCGGCTTGGGAATTAAGTTTG	70535
Qy	2852	TTTCTGGCCTTTAGCCTA	2869
Db	70534	TTTCTGGCCTTTAGCCTA	70517

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 06:15:30 ; Search time 1086.48 Seconds
(without alignments)
11511.184 Million cell updates/sec

Title: US-10-669-693-1

Perfect score: 2944

Sequence: 1 tttctctgttgcttactc.....aaaaaaaaaaaaaaaaaaaa 2944

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2000s:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2944	100.0	2944	6	AAD45438
2	2621.4	89.0	2623	4	AAS21275 Human CDN
3	2621.4	89.0	2623	7	ACD23884 Novel hum
4	2621.4	89.0	2623	7	ACA67025 cDNA enco
5	2621.4	89.0	2623	7	ACA03634 cDNA enco
6	2621.4	89.0	2623	7	ABX89172 DNA enco
7	2621.4	89.0	2623	7	ACD41826 Human sec
8	2621.4	89.0	2623	7	ACA04055 Human CDN
9	2621.4	89.0	2623	8	ADA45582 Novel hum
10	2621.4	89.0	2623	8	ADA76013 Human PRO
11	2621.4	89.0	2623	8	ADA18663 Human PRO
12	2621.4	89.0	2623	8	ADA61286 Homo sapi
13	2621.4	89.0	2623	8	ADB19071 Novel hum
14	2621.4	89.0	2623	8	ADB27612 cDNA enco
15	2621.4	89.0	2623	8	ADA86091 Novel hum
16	2621.4	89.0	2623	8	ADB15655 Human PRO
17	2621.4	89.0	2623	8	ADA47441 Human PRO
18	2621.4	89.0	2623	8	ADA67236 Human PRO
19	2621.4	89.0	2623	8	ADB30243 cDNA enco
20	2621.4	89.0	2623	8	ADA85539 Novel hum
21	2621.4	89.0	2623	8	ADA96751 Human PRO
22	2621.4	89.0	2623	8	ADA79055 Human PRO
23	2621.4	89.0	2623	8	ADA87194 Novel hum

24	2621.4	89.0	2623	8	ADB16396	Human PRO
25	2621.4	89.0	2623	8	ADA91488	Novel hum
26	2621.4	89.0	2623	8	ADB14551	Human PRO
27	2621.4	89.0	2623	8	ADB18512	Novel hum
28	2621.4	89.0	2623	8	ADA93727	Human PRO
29	2621.4	89.0	2623	8	ADB19623	Novel hum
30	2621.4	89.0	2623	8	ADB12935	Human PRO
31	2621.4	89.0	2623	8	ACD98455	Novel hum
32	2621.4	89.0	2623	8	ADA74189	Human PRO
33	2621.4	89.0	2623	8	ADB24422	Human PRO
34	2621.4	89.0	2623	8	ADA81946	Human PRO
35	2621.4	89.0	2623	8	ADA74909	Human PRO
36	2621.4	89.0	2623	8	ADA84987	Novel hum
37	2621.4	89.0	2623	8	ADA84435	Novel hum
38	2621.4	89.0	2623	8	ADB29691	cDNA enco
39	2621.4	89.0	2623	8	ADA80219	Human PRO
40	2621.4	89.0	2623	8	ADA75461	Human PRO
41	2621.4	89.0	2623	8	ADA46686	Human PRO
42	2621.4	89.0	2623	8	ADB24982	Human PRO
43	2621.4	89.0	2623	8	ADA93158	Human PRO
44	2621.4	89.0	2623	8	ADB26508	cDNA enco
45	2621.4	89.0	2623	8	ADB30795	cDNA enco

ALIGNMENTS

RESULT 1
AAD45438
ID AAD45438 standard; cDNA; 2944 BP.
XX
AC AAD45438;
XX
DT 27-DEC-2002 (first entry)
XX
DE Human drug-metabolising enzyme encoding cDNA.
XX
KW Human; drug-metabolising enzyme; cytochrome p450; therapy; cytostatic;
KW Parkinson's disease; myotonic dystrophy; developmental defect; enzyme;
KW cancer; chromosome 19; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..303
FT FT /*tag= a
FT CDS 304..1818
FT FT /*tag= b
FT FT /product= "Human drug-metabolising enzyme"
FT 3'UTR 1819..2944
FT FT /*tag= c
XX
US2002076774-A1.
XX
20-JUN-2002.
XX
27-DEC-2000; 2000US-00748127.
XX
21-JUN-2000; 2000US-0212840P.
XX
(YANC/) YAN C.
XX (SHAO/) SHAO W.
XX (MERK/) MERKULOV G V.
XX (DFRA/) DI FRANCESCO V.
XX (BEAS/) BEASLEY E M.
XX
PI Yan C, Shao W, Merkulov GV, Di Francesco V, Beasley EM;
XX
WPI; 2002-673347/72.
XX P-PSDB; AAE28282.
XX
PT A new human drug-metabolizing enzyme is related to the cytochrome p450
PT superfamily and is useful to provide modulators to treat drug-

PT metabolizing enzyme-related disorders including cancer and Parkinson's
 PT disease.
 XX
 PS
 PS
 XX Claim 4; Page 20-22; 49pp; English.
 XX
 CC The invention relates to human drug-metabolising enzyme peptides related
 CC to the cytochrome p450 superfamily. Binding agents of the peptide are
 CC used to treat a disease or condition mediated by a human drug-
 CC metabolising enzyme. Such diseases include Parkinson's disease, cancer,
 CC myotonic dystrophy and developmental defects. The present sequence is
 CC human drug-metabolising enzyme encoding cDNA. The gene encoding this
 CC enzyme is located at chromosome 19
 XX
 XX Sequence 2944 BP; 648 A; 957 C; 794 G; 545 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 2944; DB 6; Length 2944;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTTCTTCTGTTGCTTACTCCCTATCCGGGGGCGCAAGGCGCTGTCGGCGGCCAAGC 60
 DB 1 TTTCTTCTGTTGCTTACTCCCTATCCGGGGGCGCAAGGCGCTGTCGGCGGCCAAGC 60
 QY 61 CCGCGGTAAACCTGGGTGACCTCGAGACATCCGTTGGAGCATGAGTTCGCCACATCAGG 120
 DB 61 CCGCGGTAAACCTGGGTGACCTCGAGACATCCGTTGGAGCATGAGTTCGCCACATCAGG 120
 QY 121 CGCGCGCGGTGTCTCGGGAGAAACCGCGGCGGGGGAGATAAGCTGCGCAGAGGACAGG 180
 DB 121 CGCGCGCGGTGTCTCGGGAGAAACCGCGGCGGGGGAGATAAGCTGCGCAGAGGACAGG 180
 QY 181 GGGCTGGGTAGTGTGCGCGCGCGCGCGCTGCTGTTGGGGAGAGACCGCGGCTC 240
 DB 181 GGGCTGGGTAGTGTGCGCGCGCGCGCGCTGCTGTTGGGGAGAGACCGCGGCTC 240
 QY 241 CGCGCCCTAACTAGCCAGCCGCGGCGGCGCTGGAGAGAGAGAGAGAGAGAGAGAGAG 300
 DB 241 CGCGCCCTAACTAGCCAGCCGCGGCGGCGCTGGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 301 GAGATGGAGCGGACCGGACCTGGGCGCTGCTGCTGGCGCTGGCGCTGCTGCTGCTG 360
 DB 301 GAGATGGAGCGGACCGGACCTGGGCGCTGCTGCTGGCGCTGGCGCTGCTGCTGCTG 360
 QY 361 AGCTGGGCGCTGTCGGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 DB 361 AGCTGGGCGCTGTCGGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 QY 421 CCACCTGCTGGGAAACCTCTGCTGAGCTACGGGCGGCGGCGGCGGCTGTATTCAGGGCTCATGCGG 480
 DB 421 CCACCTGCTGGGAAACCTCTGCTGAGCTACGGGCGGCGGCGGCGGCTGTATTCAGGGCTCATGCGG 480
 QY 481 CTGAGTAAGAGTAACGAGCGGTGTTTCACTACCTAGCTGCGGACCTGGGCGGCGGCTG 540
 DB 481 CTGAGTAAGAGTAACGAGCGGTGTTTCACTACCTAGCTGCGGACCTGGGCGGCGGCTG 540
 QY 541 GTCTGTTGGGAGAGAGCTGCGGGAGGCGGCTGGGAGGCTGAGGCTGAGGAGGTTTTC 600
 DB 541 GTCTGTTGGGAGAGAGCTGCGGGAGGCGGCTGGGAGGCTGAGGCTGAGGAGGTTTTC 600
 QY 601 GGCCTGGGAAACCGTAGCGATGCTGGGAGGAGCTTTTGTATGGCCATGGGGTTTCTCTCC 660
 DB 601 GGCCTGGGAAACCGTAGCGATGCTGGGAGGAGCTTTTGTATGGCCATGGGGTTTCTCTCC 660
 QY 661 AACGGGAGCGGTGGAGGAGCTGAGGAGGTTTACCATGCTGCTGCGGACCTGGGCG 720
 DB 661 AACGGGAGCGGTGGAGGAGCTGAGGAGGTTTACCATGCTGCTGCGGACCTGGGCG 720
 QY 721 ATGGGGAAGCGAAGCGAGGAGCTGATCCAGCGGAGCGCGCGGTGCTGTTGGAGACA 780
 DB 721 ATGGGGAAGCGAAGCGAGGAGCTGATCCAGCGGAGCTGATCCAGCGGAGCGCGGTG 780
 QY 781 TTCAGGGGACAGAGGACCGCCATTGATTCCTTCCCTGCTGCTGCGCCAGGCGACCTCC 840
 DB 781 TTCAGGGGACAGAGGACCGCCATTGATTCCTTCCCTGCTGCTGCGCCAGGCGACCTCC 840

DB 781 TTCAGGGGACAGAGGACCGCCATTGATTCCTTCCCTGCTGCTGCGCCAGGCGACCTCC 840
 QY 841 AACGTAGTCTGCTCCTCCTCTTTGGGCTCGGCTTCTCTATGAGATAAGAGTTCAG 900
 DB 841 AACGTAGTCTGCTCCTCCTCTTTGGGCTCGGCTTCTCTATGAGATAAGAGTTCAG 900
 QY 901 GCCGTGGTCCGGGACAGCTGGTGTACCTCTGCTGGAGTCAGCTCCAGGGGGTTCAGACC 960
 DB 901 GCCGTGGTCCGGGACAGCTGGTGTACCTCTGCTGGAGTCAGCTCCAGGGGGTTCAGACC 960
 QY 961 TACGAGATGTTCTCCTGTTCTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 DB 961 TACGAGATGTTCTCCTGTTCTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 QY 1021 CACGTGACGACCTTGGCTGCTTACAGTCCGGGAGGTGCAGCAGCAGCAGGGAACCTG 1080
 DB 1021 CACGTGACGACCTTGGCTGCTTACAGTCCGGGAGGTGCAGCAGCAGCAGGGAACCTG 1080
 QY 1081 GATGCTTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 DB 1081 GATGCTTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 QY 1141 GAACAAAACCCAGGACAGAAATTCACCAAGAAACATGCTGATGACAGTCAATTTATTG 1200
 DB 1141 GAACAAAACCCAGGACAGAAATTCACCAAGAAACATGCTGATGACAGTCAATTTATTG 1200
 QY 1201 CTGTTTGTGGGACGATGACGCTGAGCAGCAGCGTGGGTATACCTCTGCTCTCTGATG 1260
 DB 1201 CTGTTTGTGGGACGATGACGCTGAGCAGCAGCGTGGGTATACCTCTGCTCTCTGATG 1260
 QY 1261 AAATACCTCTCATGTCCAAAAGTGGGTAGTGGAGAGTGAATCGGGAGTGGGGGCTGCG 1320
 DB 1261 AAATACCTCTCATGTCCAAAAGTGGGTAGTGGAGAGTGAATCGGGAGTGGGGGCTGCG 1320
 QY 1321 CAGGACCAAGGCTTGGGAGCGGTACCGGCTTCCCTTACACGAGCGGCTTCTGATGAG 1380
 DB 1321 CAGGACCAAGGCTTGGGAGCGGTACCGGCTTCCCTTACACGAGCGGCTTCTGATGAG 1380
 QY 1381 GCGGAGCGGCTGCTGGCGTGGTCCCATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
 DB 1381 GCGGAGCGGCTGCTGGCGTGGTCCCATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
 QY 1441 CGCTTCCGAGGCTACACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1500
 DB 1441 CGCTTCCGAGGCTACACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1500
 QY 1501 CTGCTATGACCCCAACATCTTCAAGCACCCAGAGAGTTCACCCAGACCGTTCCTGGAT 1560
 DB 1501 CTGCTATGACCCCAACATCTTCAAGCACCCAGAGAGTTCACCCAGACCGTTCCTGGAT 1560
 QY 1561 GCAGATGACCGTTCAGGAGCATGAGGCGTTCCTGCGGCTTCTCTTACCCAGGAGCGGTC 1620
 DB 1561 GCAGATGACCGTTCAGGAGCATGAGGCGTTCCTGCGGCTTCTCTTACCCAGGAGCGGTC 1620
 QY 1621 TGCTTGGAGAGGCGCTTGGGAGGAGGAGTTCCTCTTCTTCTTACCCAGGAGCGGTC 1680
 DB 1621 TGCTTGGAGAGGCGCTTGGGAGGAGGAGTTCCTCTTCTTCTTACCCAGGAGCGGTC 1680
 QY 1681 GCGTCTTCTGAGAGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1740
 DB 1681 GCGTCTTCTGAGAGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1740
 QY 1741 GGCCTTTTCAACATTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1800
 DB 1741 GGCCTTTTCAACATTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1800
 QY 1801 ACCAGCAGACAGATGAGAGGAGGCACTTGGAGTGGTGGTGGTGGTGGTGGTGGTGG 1860
 DB 1801 ACCAGCAGACAGATGAGAGGAGGCACTTGGAGTGGTGGTGGTGGTGGTGGTGGTGG 1860
 QY 1861 CCAGCTCAACAGTGGGATGGAGGAGGTTAATGCTCCAGAGTGTACACTGCGAGGCGAG 1920
 DB 1861 CCAGCTCAACAGTGGGATGGAGGAGGTTAATGCTCCAGAGTGTACACTGCGAGGCGAG 1920

1921 QY CACATTTACACGGCTGCAGTTGTTTTCCGGAGTCTGTCCACGGCCACACGCTCACTTG 1980
1921 Db CACATTTACACGGCTGCAGTTGTTTTCCGGAGTCTGTCCACGGCCACACGCTCACTTG 1980
1981 QY ACTCATGCTGCTAAGATGCAACCGCACACCCATACACAACTACAAGGGCCACAAAGCA 2040
1981 Db ACTCATGCTGCTAAGATGCAACCGCACACCCATACACAACTACAAGGGCCACAAAGCA 2040
2041 QY ACTGCTGGGTAGTTTCCACAGACATAAATATAGTCCATCTGCAATCAAGCACATAG 2100
2041 Db ACTGCTGGGTAGTTTCCACAGACATAAATATAGTCCATCTGCAATCAAGCACATAG 2100
2101 QY CCAAGTAACCCCAACTCCCTCGATCTGCAGCCACACGTGGAGTCTGGGTGTCACC 2160
2101 Db CCAAGTAACCCCAACTCCCTCGATCTGCAGCCACACGTGGAGTCTGGGTGTCACC 2160
2161 QY TTCACAAGCCACAGAAACGGCCACACATGTTTCACAGCTCACAGCCCTCTCCCATTCATCG 2220
2161 Db TTCACAAGCCACAGAAACGGCCACACATGTTTCACAGCTCACAGCCCTCTCCCATTCATCG 2220
2221 QY AACTTCTCAGTGTCCCTGTCTGCTGCTGGCCACAGGGAACAGCATGCCCCCTCCGGGG 2280
2221 Db AACTTCTCAGTGTCCCTGTCTGCTGCTGGCCACAGGGAACAGCATGCCCCCTCCGGGG 2280
2281 QY TCATGCCACCCAGAGACTGTCGCTGCTATGCGCCCAACTCATGCTCCCTCTTGGCTA 2340
2281 Db TCATGCCACCCAGAGACTGTCGCTGCTATGCGCCCAACTCATGCTCCCTCTTGGCTA 2340
2341 QY CACCACTCTCCAGCCTGTGACCAACCGATGTCACACACACCCCAACCACTTGTCCACACA 2400
2341 Db CACCACTCTCCAGCCTGTGACCAACCGATGTCACACACACCCCAACCACTTGTCCACACA 2400
2401 QY GCTTACCCAGTACGACATGCTCTGCTGCCCTCCAGAGATATCTCCCACTGAGACACGCCGC 2460
2401 Db GCTTACCCAGTACGACATGCTCTGCTGCCCTCCAGAGATATCTCCCACTGAGACACGCCGC 2460
2461 QY CCCCACAGAGGCACAGTCCCGACGCACTCTGCAACTGCAAGCCCTCAGTCAACCCCTTTT 2520
2461 Db CCCCACAGAGGCACAGTCCCGACGCACTCTGCAACTGCAAGCCCTCAGTCAACCCCTTTT 2520
2521 QY AAGCACCTTGATTTCTACCAATGCAACACATCTGGGTCTGCGATTATGACACAGACACT 2580
2521 Db AAGCACCTTGATTTCTACCAATGCAACACATCTGGGTCTGCGATTATGACACAGACACT 2580
2581 QY TGAACATACAGGACCTCTGACCGGAGGAACCTTGCCCAACCCCAACACGTCGTTATG 2640
2581 Db TGAACATACAGGACCTCTGACCGGAGGAACCTTGCCCAACCCCAACACGTCGTTATG 2640
2641 QY TAACACGTTGGAAGCGCCCTGCTGCCCTCCACACACACATACACTCACTGATCT 2700
2641 Db TAACACGTTGGAAGCGCCCTGCTGCCCTCCACACACACATACACTCACTGATCT 2700
2701 QY ACAGCCCTGTTCCGGCTCAGAGTCCCACTAGACCCAGTGAAGGGTTAGAGACCAAG 2760
2701 Db ACAGCCCTGTTCCGGCTCAGAGTCCCACTAGACCCAGTGAAGGGTTAGAGACCAAG 2760
2761 QY TAGGGCCAGTTTCCAAATTCACCTGTACGGAGTGAAGCGGATCTGACGTTCTTGTGA 2820
2761 Db TAGGGCCAGTTTCCAAATTCACCTGTACGGAGTGAAGCGGATCTGACGTTCTTGTGA 2820
2821 QY CTTTAAAGGTCGGCTTGGGAATTAAGTTTGTCTGGCCTTTAGCCTAAAAA 2880
2821 Db CTTTAAAGGTCGGCTTGGGAATTAAGTTTGTCTGGCCTTTAGCCTAAAAA 2880
2881 QY AA 2940
2881 Db AA 2940
2941 QY AAAAA 2944
2941 Db AAAAA 2944

RESULT 2
AAS21275
ID AAS21275 standard; cDNA; 2623 BP.

XX

AC AAS21275;

DT 24-OCT-2001 (first entry)

XX

Human cDNA sequence encoding for PRO1906 polypeptide.

Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;
prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;
ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;
A-peptide; factor VIIA; gene therapy; ss.

OS Homo sapiens.

XX WO200140466-A2.

XX

PD 07-JUN-2001.

XX 01-DEC-2000; 2000WO-US032678.

XX 01-DEC-1999; 99WO-US028301.

XX 01-DEC-1999; 99WO-US028634.

XX 02-DEC-1999; 99WO-US028551.

XX 02-DEC-1999; 99WO-US028564.

XX 02-DEC-1999; 99WO-US028565.

XX 09-DEC-1999; 99US-0170262P.

XX 16-DEC-1999; 99WO-US030095.

XX 20-DEC-1999; 99WO-US030911.

XX 30-DEC-1999; 99WO-US031243.

XX 30-DEC-1999; 99WO-US031274.

XX 05-JAN-2000; 2000WO-US000219.

XX 06-JAN-2000; 2000WO-US000277.

XX 06-JAN-2000; 2000WO-US000376.

XX 11-FEB-2000; 2000WO-US003565.

XX 18-FEB-2000; 2000WO-US004341.

XX 22-FEB-2000; 2000WO-US004342.

XX 24-FEB-2000; 2000WO-US004914.

XX 24-FEB-2000; 2000WO-US005004.

XX 01-MAR-2000; 2000WO-US005601.

XX 02-MAR-2000; 2000US-0187202P.

XX 10-MAR-2000; 2000WO-US006319.

XX 15-MAR-2000; 2000WO-US006884.

XX 20-MAR-2000; 2000WO-US007377.

XX 21-MAR-2000; 2000WO-US007532.

XX 30-MAR-2000; 2000WO-US008439.

XX 17-MAY-2000; 2000WO-US013705.

XX 22-MAY-2000; 2000WO-US014042.

XX 30-MAY-2000; 2000WO-US014941.

XX 02-JUN-2000; 2000WO-US015264.

XX 03-JUN-2000; 2000US-0209832P.

XX 28-JUL-2000; 2000WO-US020710.

XX 11-AUG-2000; 2000WO-US020231.

XX 23-AUG-2000; 2000WO-US023522.

XX 24-AUG-2000; 2000WO-US023328.

XX 08-NOV-2000; 2000WO-US030952.

XX 10-NOV-2000; 2000WO-US030873.

(GETH) GENENTECH INC.

PA

XX

PI

PI

PI

XX

DR

XX

DR

XX

XX

XX

Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
Geritsen ME, Goddard A, Godowski PU, Gurney AL, Sherwood S;
Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
WPI; 2001-408281/43.
P-PSDB; AAU12203.

QY 1944 TTTCGGAGTCTCTCCACAGCCCCACACAGCTCACTTGACTCATGCTGCTAAGATGCACAA 2003
DB |||
QY 1981 TTTCGGAGTCTGTCACAGCCCCACACAGCTCACTTGACTCATGCTGCTAAGATGCACAA 1740
DB |||
QY 2004 CCGCACACCCATACACAACTACAAGGGCCACAAAGCAACTCTCTGGTTAGCTTTCCACAG 2063
DB |||
QY 1741 CCGCACACCCATACACAACTACAAGGGCCACAAAGCAACTCTCTGGTTAGCTTTCCACAG 1800
DB |||
QY 2064 ACATAAATATAGTCCATCTGCAATCACAGACATAGCCAGGTAAACCCACCACTCCGCT 2123
DB |||
QY 1801 ACATAAATATAGTCCATCTGCAATCACAGACATAGCCAGGTAAACCCACCACTCCGCT 1860
DB |||
QY 2124 GGATCTGCAGCCACACAGTGGAGTCTGGCTGTCACTTCAAGCCACACAGAAAGGCCA 2183
DB |||
QY 1861 GGATCTGCAGCCACACAGTGGAGTCTGGCTGTCACTTCAAGCCACACAGAAAGGCCA 1920
DB |||
QY 2184 CACATGTTCAAGTCTACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 2243
DB |||
QY 1921 CACATGTTCAAGTCTACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 1980
DB |||
QY 2244 GGTGCTGGCACAGGGAACAGTATGCCCTCCGGGTCTATGCCACCCAGAGACTGTGCG 2303
DB |||
QY 1981 GGTGCTGGCACAGGGAACAGTATGCCCTCCGGGTCTATGCCACCCAGAGACTGTGCG 2040
DB |||
QY 2304 TGTCTATGGCCCAACTCATGCTCCCTCTCTTGCTACACCACTCTCCAGCCGTGAC 2363
DB |||
QY 2041 TGTCTATGGCCCAACTCATGCTCCCTCTCTTGCTACACCACTCTCCAGCCGTGAC 2100
DB |||
QY 2364 ACCGATGTCCACACACCCCAACCACTTGCTGCCACAGCTACCCAGCTAGACATCGTCC 2423
DB |||
QY 2101 ACCGATGTCCACACACCCCAACCACTTGCTGCCACAGCTACCCAGCTAGACATCGTCC 2160
DB |||
QY 2424 TGGTCTCCACAGATGATCTTCCACTGAGACACGCGGCCCCACAGAGGCACAGTCCCGAG 2483
DB |||
QY 2161 TGGTCTCCACAGATGATCTTCCACTGAGACACGCGGCCCCACAGAGGCACAGTCCCGAG 2220
DB |||
QY 2484 CCACCTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTACCAATG 2543
DB |||
QY 2221 CCACCTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTACCAATG 2280
DB |||
QY 2544 CAAACACATCTGGTCTGCGATTATGCACAGAGACTTTGGACATACAGGACCCCTCAGAC 2603
DB |||
QY 2281 CAAACACATCTGGTCTGCGATTATGCACAGAGACTTTGGACATACAGGACCCCTCAGAC 2340
DB |||
QY 2604 CGGAGGAACCTGCCCAACCCCAACACGCTGCTTATGTAACCACTGGAAGCGGCCCT 2663
DB |||
QY 2341 CGGAGGAACCTGCCCAACCCCAACACGCTGCTTATGTAACCACTGGAAGCGGCCCT 2400
DB |||
QY 2664 GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGGCGTCAAG 2723
DB |||
QY 2401 GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGGCGTCAAG 2460
DB |||
QY 2724 TCCCCACTAGACCCAGTGAAGGGTTAGAGCAAGTAGGGCCAGTTTCCAATTCCACC 2783
DB |||
QY 2461 TCCCCACTAGACCCAGTGAAGGGTTAGAGCAAGTAGGGCCAGTTTCCAATTCCACC 2520
DB |||
QY 2784 CTGTCAGGAGTGAAGCGGATCTGACGCTCTGTTGACTTAAGGTCGGCTTGGGAATT 2843
DB |||
QY 2521 CTGTCAGGAGTGAAGCGGATCTGACGCTCTGTTGACTTAAGGTCGGCTTGGGAATT 2580
DB |||
QY 2844 AAAGTTTGTCTTGGCCCTTTAGCCTAAAAAAGGAAAAA 2886
DB |||
QY 2581 AAAGTTTGTCTTGGCCCTTTAGCCTAAAAAAGGAAAAA 2623
DB |||

RESULT 3

ACD23884

ID ACD23884 standard; cDNA; 2623 BP.

XX

AC ACD23884;

XX

DT 26-AUG-2003 (first entry)

XX

DE Novel human secreted and transmembrane protein PRO1906 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; antiinflammatory;
KW antiarteriosclerotic; cardiant; anti-infertility; anti-HIV; cytostatic;
KW antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;
KW TNF-alpha release; cell proliferation; cell differentiation;
KW gene expression modulator; proteoglycan release; cytokine release;
KW tumour; inflammatory disease; organ failure; atherosclerosis;
KW cardiac injury; infertility; birth defect; premature aging; AIDS;
KW acquired immunodeficiency syndrome; cancer; diabetic complication;
KW chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;
KW bioreactor; tissue typing; gene; ss.

Homo sapiens.

US2003032156-A1.

13-FEB-2003.

06-MAY-2002; 2002US-00140474.

31-MAR-1997; 97WO-US005230.

12-JUN-1998; 98WO-US012456.

14-JUL-1998; 98WO-US014552.

28-AUG-1998; 98WO-US017888.

10-SEP-1998; 98WO-US018824.

14-SEP-1998; 98WO-US019093.

14-SEP-1998; 98WO-US019094.

16-SEP-1998; 98WO-US019177.

17-SEP-1998; 98WO-US019437.

07-OCT-1998; 98WO-US021141.

29-OCT-1998; 98WO-US022991.

29-OCT-1998; 98WO-US022992.

20-NOV-1998; 98WO-US024855.

01-DEC-1998; 98WO-US025108.

05-JAN-1999; 99WO-US000106.

08-MAR-1999; 99WO-US0005028.

10-MAR-1999; 99WO-US0005190.

20-APR-1999; 99WO-US008615.

14-MAY-1999; 99WO-US010733.

02-JUN-1999; 99WO-US012252.

01-SEP-1999; 99WO-US020111.

08-SEP-1999; 99WO-US020594.

13-SEP-1999; 99WO-US020944.

15-SEP-1999; 99WO-US021090.

15-SEP-1999; 99WO-US021547.

05-OCT-1999; 99WO-US023089.

29-NOV-1999; 99WO-US028214.

30-NOV-1999; 99WO-US028313.

01-DEC-1999; 99WO-US028409.

01-DEC-1999; 99WO-US028301.

01-DEC-1999; 99WO-US028634.

02-DEC-1999; 99WO-US028551.

02-DEC-1999; 99WO-US028564.

02-DEC-1999; 99WO-US028565.

16-DEC-1999; 99WO-US030095.

20-DEC-1999; 99WO-US030911.

20-DEC-1999; 99WO-US030999.

22-DEC-1999; 99WO-US030720.

30-DEC-1999; 99WO-US031243.

30-DEC-1999; 99WO-US031274.

03-JAN-2000; 2000WO-US000219.

06-JAN-2000; 2000WO-US000277.

06-JAN-2000; 2000WO-US000376.

11-FEB-2000; 2000WO-US003565.

18-FEB-2000; 2000WO-US004341.

18-FEB-2000; 2000WO-US004342.

22-FEB-2000; 2000WO-US004414.

24-FEB-2000; 2000WO-US004914.

24-FEB-2000; 2000WO-US005004.

01-MAR-2000; 2000WO-US005601.

02-MAR-2000; 2000WO-US005746.

PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.

XX (GETH) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2003-341980/32.
DR P-PSDB; AB017647.

XX New secreted and transmembrane PRO nucleic acids, for treating
PT inflammation, organ failure, atherosclerosis, cardiac injury,
PT infertility, birth defects, premature aging, axquired immunodeficiency
PT syndrome (AIDS), or cancer.
XX
XX Claim 2; Fig 63; 660pp; English.

XX The invention describes an isolated nucleic acid (I) comprising, or which
CC has 80 % sequence identity to, or the full-length coding sequence of, one
CC of 275 nucleotide sequences, and which encodes a corresponding
CC polypeptide selected from 275 amino acid sequences, where all sequences
CC are given in the specification. The polypeptide encoded by (I) is used to
CC detect PRO polypeptides, link a bioactive molecule to a cell expressing a
CC PRO polypeptide, modulate a biological activity of a cell, stimulate the
CC release of tumour necrosis factor (TNF)-alpha from human blood, modulate
CC the uptake of glucose or free fatty acid by cells, stimulate or inhibit

CC the proliferation or differentiation of cells or gene expression,
CC stimulate the release of proteoglycans, stimulate the release of cytokine
CC from peripheral blood mononuclear cells, inhibit the binding of A-peptide
CC to factor VIIA, or detect the presence of tumour in a mammal. The nucleic
CC acid and polypeptide encoded by it, are useful for treating inflammatory
CC diseases, organ failure, atherosclerosis, cardiac injury, infertility,
CC birth defects, premature aging, acquired immunodeficiency syndrome
CC (AIDS), cancer, or diabetic complications. The nucleic acid is useful as
CC hybridisation probes, in chromosome and gene mapping, and in generating
CC antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,
CC diagnostics, biosensors or bioreactors. Both are useful in tissue typing.
CC This sequence encodes a novel human secreted and transmembrane PRO
CC polypeptide

XX Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Query Match 89.0%; Score 2621.4; DB 7; Length 2623;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 264 GCGGAGCGCTCGGAGAGGAGGAGCGACCTGCCGAGATGGAGCGCACCGGACCTTG 323
Db 1 GCGGAGCGCTCGGAGAGGAGGAGCGACCTGCCGAGATGGAGCGCACCGGACCTTG 60
QY 324 GCGGCTGTGCTGGCGCTGGCGCTGCTCTGCTGTGACGCTGGCGCTCTCGGGACCA 383
Db 61 GCGGCTGTGCTGGCGCTGGCGCTGCTCTGCTGTGACGCTGGCGCTCTCGGGACCA 120
QY 384 GCGCGAGGCGACCTGCCCCCGGGCCACGGCGGTACACTGCTGGGAAACCTCTGCA 443
Db 121 GCGCGAGGCGACCTGCCCCCGGGCCACGGCGGTACACTGCTGGGAAACCTCTGCA 180
QY 444 GCTACGGCCCCGGCGCTGTATTTCAGGGCTCATGGCGGTGAGTAAAGTACGACCGGT 503
Db 181 GCTACGGCCCCGGCGCTGTATTTCAGGGCTCATGGCGGTGAGTAAAGTACGACCGGT 240
QY 504 GTTCACCATCTACCTGGGACCTGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 563
Db 241 GTTCACCATCTACCTGGGACCTGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 564 GCGGAGGCGCTGGGAGGTGAGGAGTTCAGGCGCGCGGGGAAACGTAGCGATGCT 623
Db 301 GCGGAGGCGCTGGGAGGTGAGGAGTTCAGGCGCGCGGGGAAACGTAGCGATGCT 360
QY 624 GGAAGGAGCTTTTGATGGCCATGGGGTTTTTCTTCTCAACGGGAGCGGTGGAGGAGCT 683
Db 361 GGAAGGAGCTTTTGATGGCCATGGGGTTTTTCTTCTCAACGGGAGCGGTGGAGGAGCT 420
QY 684 GAGGAAGTTTACATGCTTGCTCTCGGACCTGGGCATGGGAAACGAGAGCGGAGGA 743
Db 421 GAGGAAGTTTACATGCTTGCTCTCGGACCTGGGCATGGGAAACGAGAGCGGAGGA 480
QY 744 GCTGATCCAGGCGGAGGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 803
Db 481 GCTGATCCAGGCGGAGGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 804 ATTTCATCCCTCCCTGCTGCTGCGCCAGGCCACCTCCAAACGTAGTGTGTCTCCCTCTCT 863
Db 541 ATTTCATCCCTCCCTGCTGCTGCGCCAGGCCACCTCCAAACGTAGTGTGTCTCCCTCTCT 600
QY 864 TGGCCTCCGCTTCTCTATGAGATAAGAGTGTTCAGGCGGTGTGTGTGTGTGTGTGTGTGT 923
Db 601 TGGCCTCCGCTTCTCTATGAGATAAGAGTGTTCAGGCGGTGTGTGTGTGTGTGTGTGTGT 660
QY 924 TACCCTGTGGGAGTCAGTCCCGGGGGTTCAGACTACGAGATGTTCTCTGTGTCTCT 983
Db 661 TACCCTGTGGGAGTCAGTCCCGGGGGTTCAGACTACGAGATGTTCTCTGTGTCTCT 720
QY 984 GCGGCGCTTCGCCAGGCGCGCCCAAGCAGTCTCTCCACACGTACGACCTTGCTGCTCTCT 1043
Db 721 GCGGCGCTTCGCCAGGCGCGCCCAAGCAGTCTCTCCACACGTACGACCTTGCTGCTCTCT 780
QY 1044 CACAGTCCGGCGGTGCAGCAGCACCGAGGGAACCTGGATGCTTCCGGGCCCCGACGTGA 1103

[illegible]

Db	1861	GGATCTGCAGCCCAACA	CGTGGGAGTCTGGCTGTCA	CCCTTCAAGGCCACAGAAACGGCCA	1920
Qy	2184	CACATGTTTCACAGCTCA	CACGCCCTCTCCATTAT	ATCGAACTTCTCAGTGTCCCTGTCCCT	2243
Db	1921	CACATGTTTCACAGCTCA	CACGCCCTCTCCATTAT	ATCGAACTTCTCAGTGTCCCTGTCCCT	1980
Qy	2244	GGTGCTGGCACAGGGAA	CAGCATGCCCCCTCGGGGT	CATGCCACCACGAGACTGTGCG	2303
Db	1981	GGTGCCTGGCACAGGGAA	CAGCATGCCCCCTCGGGGT	CATGCCACCACGAGACTGTGCG	2040
Qy	2304	TGTCATATGGCCCCAACT	CACTGCTCCCTCTCTTGGCTA	CACCACTCTCCCCAGCCTGTGACC	2363
Db	2041	TGTCATATGGCCCCAACT	CACTGCTCCCTCTCTTGGCTA	CACCACTCTCCCCAGCCTGTGACC	2100
Qy	2364	ACCGATGTCACACACAC	CCCCCAACCACTGTGCCACAGCT	TACCGACGTACGACATGTGTC	2423
Db	2101	ACCGATGTCACACACAC	CCCCCAACCACTGTGCCACAGCT	TACCGACGTACGACATGTGTC	2160
Qy	2424	TGGCTCCCCAGAGTATCT	TCCCACTGAGACA	CGCCGCCCCACAGAGGCACAGTCCCCAG	2483
Db	2161	TGGCTCCCCAGAGTATCT	TCCCACTGAGACA	CGCCGCCCCACAGAGGCACAGTCCCCAG	2220
Qy	2484	CCACTCTGCAACTGACG	CCCTCAGTCA	CCCCCTTTTAAAGCACCCCTGATTCTACCAAATG	2543
Db	2221	CCACTCTGCAACTGACG	CCCTCAGTCA	CCCCCTTTTAAAGCACCCCTGATTCTACCAAATG	2280
Qy	2544	CAAAACATCTGGGTCTG	CGCAATTATGCACAGACACT	TTTGACATACAGAGACCTCAGAC	2603
Db	2281	CAAAACATCTGGGTCTG	CGCAATTATGCACAGACACT	TTTGACATACAGAGACCTCAGAC	2340
Qy	2604	CGGAGGAACACTGCCCCA	ACCCCAACAACGCTGCTTAT	GTAAACACGCTGGAAAGCGGCCCT	2663
Db	2341	CGGAGGAACACTGCCCCA	ACCCCAACAACGCTGCTTAT	GTAAACACGCTGGAAAGCGGCCCT	2400
Qy	2664	GCTGCCCTCCACACAC	ACATACACACTCACTGATCT	CACAGCCCTGTTGGCGGTGACAG	2723
Db	2401	GCTGCCCTCCACACAC	ACATACACACTCACTGATCT	CACAGCCCTGTTGGCGGTGACAG	2460
Qy	2724	TCCCACTAGACCCAGTGG	AAGGGTTAGAGACCAAGTAGGGG	CCAGTTTCCAAATTCACC	2783
Db	2461	TCCCACTAGACCCAGTGG	AAGGGTTAGAGACCAAGTAGGGG	CCAGTTTCCAAATTCACC	2520
Qy	2784	CTGTGAGGAGTGAAGCGG	ATCTGACGTTCCCTTGTGAC	TTAAGGGTCCGGCTTGGGAAT	2843
Db	2521	CTGTGAGGAGTGAAGCGG	ATCTGACGTTCCCTTGTGAC	TTAAGGGTCCGGCTTGGGAAT	2580
Qy	2844	AAAGTTGTTTCTGGCTTT	AGCCTTAAAGGCTTAAAGGCTT	GGGAAT	2886
Db	2581	AAAGTTGTTTCTGGCTTT	AGCCTTAAAGGCTTAAAGGCTT	GGGAAT	2623
RESULT 4					
ACA67025					
ID	ACA67025 standard; cdna; 2623 BP.				
XX	ACA67025;				
AC	ACA67025;				
XX	23-JUN-2003 (first entry)				
DT	cDNA encoding human PRO polypeptide #32.				
XX	Human; PRO polypeptide; secreted and transmembrane protein;				
KW	anti-PRO antibody; diagnostic assay; gene expression; diabetes;				
KW	bone disorder; cartilage disorder; rheumatoid arthritis; obesity;				
KW	sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;				
KW	hearing loss; coagulation disorder; stroke; heart attack; cardiac;				
KW	antidiabetic; anorectic; vulnerable; antiarthritic; osteopathic;				
KW	anti-rheumatic; auditory; cerebroprotective; angiogenic; gene; ss.				
OS	Homo sapiens.				
XX	US2003004311-A1.				
PN	XX				
XX	XX				

QY	2544	CAAAACATCTGGCTCTGGATTATGCACAGAGACTTTGGACATACGAGGACCTTCAGAC	2603	PR	15-SEP-1999;	99WO-US021090.
DB	2281	CAAAACATCTGGGTCGGATTATGCACAGAGACTTTGGACATACGAGGACCTTCAGAC	2340	PR	15-SEP-1999;	99WO-US021547.
QY	2604	CGAGGAACACCTGCGCCAAACCCCAACACGTGCTTATGTAACACAGTGAAGCGGCCCT	2663	PR	05-OCT-1999;	99WO-US023089.
DB	2341	CGAGGAACACCTGCGCCAAACCCCAACACGTGCTTATGTAACACAGTGAAGCGGCCCT	2400	PR	29-NOV-1999;	99WO-US028214.
QY	2664	GTCGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTCCGCGTCAGAG	2723	PR	30-NOV-1999;	99WO-US028313.
DB	2401	GTCGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTCCGCGTCAGAG	2460	PR	30-NOV-1999;	99WO-US028409.
QY	2724	TCCCACTAGACCCAGTGAAGGGTTAGAGCCCAAGTAGGGCCAGTTTCCCAATTCACC	2783	PR	01-DEC-1999;	99WO-US028301.
DB	2461	TCCCACTAGACCCAGTGAAGGGTTAGAGCCCAAGTAGGGCCAGTTTCCCAATTCACC	2520	PR	01-DEC-1999;	99WO-US028634.
QY	2784	CTGTGAGGAGTGAGCCGGATCTGACGTTCTCTTGACCTTAAGGGTCCGGCTTGGGAATT	2843	PR	02-DEC-1999;	99WO-US028551.
DB	2521	CTGTGAGGAGTGAGCCGGATCTGACGTTCTCTTGACCTTAAGGGTCCGGCTTGGGAATT	2580	PR	02-DEC-1999;	99WO-US028564.
QY	2844	AAAGTTTGTCTGGCTTTAGCCTTAAACCAAAAAAAAAAAAAAAAAAAAA	2886	PR	02-DEC-1999;	99WO-US028565.
DB	2581	AAAGTTTGTCTGGCTTTAGCCTTAAACCAAAAAAAAAAAAAAAAAAAAA	2623	PR	16-DEC-1999;	99WO-US030095.
RESULT 5				PR	20-DEC-1999;	99WO-US030911.
ACAO3634				PR	20-DEC-1999;	99WO-US030999.
XX				PR	22-DEC-1999;	99WO-US030720.
AC				PR	30-DEC-1999;	99WO-US031243.
XX				PR	30-DEC-1999;	99WO-US031274.
DT				PR	05-JAN-2000;	2000WO-US000219.
XX				PR	06-JAN-2000;	2000WO-US000277.
DE				PR	06-JAN-2000;	2000WO-US000376.
KW				PR	11-FEB-2000;	2000WO-US000365.
KW				PR	18-FEB-2000;	2000WO-US0004341.
KW				PR	18-FEB-2000;	2000WO-US0004342.
OS				PR	22-FEB-2000;	2000WO-US000414.
PN				PR	24-FEB-2000;	2000WO-US0004914.
XX				PR	24-FEB-2000;	2000WO-US0005004.
XX				PR	01-MAR-2000;	2000WO-US0005601.
XX				PR	02-MAR-2000;	2000WO-US0005746.
XX				PR	02-MAR-2000;	2000WO-US0005841.
XX				PR	10-MAR-2000;	2000WO-US0006319.
XX				PR	15-MAR-2000;	2000WO-US0006884.
XX				PR	20-MAR-2000;	2000WO-US0007377.
XX				PR	21-MAR-2000;	2000WO-US0007532.
XX				PR	30-MAR-2000;	2000WO-US0008439.
XX				PR	17-MAY-2000;	2000WO-US013705.
XX				PR	22-MAY-2000;	2000WO-US014042.
XX				PR	30-MAY-2000;	2000WO-US014941.
XX				PR	02-JUN-2000;	2000WO-US015264.
XX				PR	28-JUL-2000;	2000WO-US020710.
XX				PR	11-AUG-2000;	2000WO-US022031.
XX				PR	23-AUG-2000;	2000WO-US023522.
XX				PR	24-AUG-2000;	2000WO-US023328.
XX				PR	08-NOV-2000;	2000WO-US030952.
XX				PR	10-NOV-2000;	2000WO-US030873.
XX				PR	01-DEC-2000;	2000WO-US032678.
XX				PR	20-DEC-2000;	2000US-00747259.
XX				PR	20-DEC-2000;	2000WO-US034956.
XX				PR	28-FEB-2001;	2001US-00796498.
XX				PR	28-FEB-2001;	2001US-00796498.
XX				PR	01-MAR-2001;	2001WO-US006666.
XX				PR	09-MAR-2001;	2001US-00802706.
XX				PR	14-MAR-2001;	2001US-00808689.
XX				PR	22-MAR-2001;	2001US-00816744.
XX				PR	05-APR-2001;	2001US-00828366.
XX				PR	10-MAY-2001;	2001US-00854208.
XX				PR	18-MAY-2001;	2001US-00854280.
XX				PR	25-MAY-2001;	2001US-00860216.
XX				PR	25-MAY-2001;	2001US-00866028.
XX				PR	25-MAY-2001;	2001US-00866034.
XX				PR	25-MAY-2001;	2001WO-US017092.
XX				PR	01-JUN-2001;	2001US-00872035.
XX				PR	05-JUN-2001;	2001US-00874503.
XX				PR	14-JUN-2001;	2001US-00882636.
XX				PR	19-JUN-2001;	2001US-00886342.
XX				PR	20-JUN-2001;	2001WO-US019692.
XX				PR	21-JUN-2001;	2001US-00887879.
XX				PR	22-JUN-2001;	2001WO-US020116.
XX				PR	29-JUN-2001;	2001WO-US021066.
XX				PR	09-JUL-2001;	2001WO-US021735.
XX				PR	18-JUL-2001;	2001US-00908827.
XX				PR	06-AUG-2001;	2001US-00924419.
XX				PR	09-AUG-2001;	2001US-00927796.

Db 541 ATTGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAAGCTAGTCTGTCTCCCTCCTCTT 600
 Qy 864 TGGCCCTCCGCTTCTCTATGAGATAAGAGTTCCAGGCCGTGTCTCCGGGACAGCTGGTG 923
 Db 601 TGGCCCTCCGCTTCTCTATGAGATAAGAGTTCCAGGCCGTGTCTCCGGGACAGCTGGTG 660
 Qy 924 TACCCTGCTGGAGTCAGCTCCAGGGGGTCCAGACCTACGAGATGTTCTCTCGGTTCT 983
 Db 661 TACCCTGCTGGAGTCAGCTCCAGGGGGTCCAGACCTACGAGATGTTCTCTCGGTTCT 720
 Qy 984 GGGGCCCTGCGAGGCCCCCAAGCAGCTCTCCACACGTCAGACCTTGGCTGCCTT 1043
 Db 721 GGGGCCCTGCGAGGCCCCCAAGCAGCTCTCCACACGTCAGACCTTGGCTGCCTT 780
 Qy 1044 CACAGTCGGCAGGTCAGCAGCACACAGGGGAACTGGATGTTGGGGCCCGCAGTGA 1103
 Db 781 CACAGTCGGCAGGTCAGCAGCACACAGGGGAACTGGATGTTGGGGCCCGCAGTGA 840
 Qy 1104 CTTTGTGATGCTTCTCTGCTGAAGATGGCAGAGGAAACAAACCCAGGCACAGAAAT 1163
 Db 841 CTTTGTGATGCTTCTCTGCTGAAGATGGCAGAGGAAACAAACCCAGGCACAGAAAT 900
 Qy 1164 CACCAACAGAACATGCTGATCAGAGTCATTTATTTGCTGTTTGTGGACGATCAGGT 1223
 Db 901 CACCAACAGAACATGCTGATCAGAGTCATTTATTTGCTGTTTGTGGACGATCAGGT 960
 Qy 1224 CAGCACCGGTCGGCTATACCTCTCTGCTGATGAATACCTCATGTCACAAAGTG 1283
 Db 961 CAGCACCGGTCGGCTATACCTCTCTGCTGATGAATACCTCATGTCACAAAGTG 1020
 Qy 1284 GGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAGGCACCAAGCTTAGGGGACG 1343
 Db 1021 GGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAGGCACCAAGCTTAGGGGACG 1080
 Qy 1344 TACCGCCTCCCTTACACGACGCGTTCTGATGAGGCGAGCGGCTGCTGGCGTGT 1403
 Db 1081 TACCGCCTCCCTTACACGACGCGTTCTGATGAGGCGAGCGGCTGCTGGCGTGT 1140
 Qy 1404 GCCCATGGGAATACCCGACACCTCATGCGGACACCGCTTCCGAGGGTACACCTGCC 1463
 Db 1141 GCCCATGGGAATACCCGACACCTCATGCGGACACCGCTTCCGAGGGTACACCTGCC 1200
 Qy 1464 CCAGGCGAGGAGTCTTCCCTCTCTTGGCTCCATCTCTGATGACCCCAACATCTCAA 1523
 Db 1201 CCAGGCGAGGAGTCTTCCCTCTCTTGGCTCCATCTCTGATGACCCCAACATCTCAA 1260
 Qy 1524 GCACCCAGAGATTCAACCCAGACGTTTCTGATGAGATGACGCGTTCCAGGAAGA 1583
 Db 1261 GCACCCAGAGATTCAACCCAGACGTTTCTGATGAGATGACGCGTTCCAGGAAGA 1320
 Qy 1584 TGAGGCGTTCCCTGCGCTTCTCTTAGGGAAGGCTGTCTGCTTGGAGGGGCTGGCAAA 1643
 Db 1321 TGAGGCGTTCCCTGCGCTTCTCTTAGGGAAGGCTGTCTGCTTGGAGGGGCTGGCAAA 1380
 Qy 1644 AGCGAGTCTTCTCTTCTTCAACCACTCTTCAAGCCTTCTCTCTGGAGGCCGCTG 1703
 Db 1381 AGCGAGTCTTCTCTTCTTCAACCACTCTTCAAGCCTTCTCTCTGGAGGCCGCTG 1440
 Qy 1704 CCCGCGGACACCTGAGCTCAAGCCACCGTCTAGTGGCTTTTCAACATTTCCCGCAGC 1763
 Db 1441 CCCGCGGACACCTGAGCTCAAGCCACCGTCTAGTGGCTTTTCAACATTTCCCGCAGC 1500
 Qy 1764 CTTCCAGTGTCAAGTCCGCTCCCTGACCTTCACTCCACGACGACAGACAGATGAAGAA 1823
 Db 1501 CTTCCAGTGTCAAGTCCGCTCCCTGACCTTCACTCCACGACGACAGATGAAGAA 1560
 Qy 1824 GCGAATTGGAGTGGTGGTCCCGAGCAGTGTCTCCAGCTCAACAGTGGGATGGA 1883
 Db 1561 GCGAATTGGAGTGGTGGTCCCGAGCAGTGTCTCCAGCTCAACAGTGGGATGGA 1620
 Qy 1884 CAGGTTAATGCTCCAGTGTACACTGACGGGACGACATTTACAGCCTGCAAGTTGT 1943
 Db 1621 CAGGTTAATGCTCCAGTGTACACTGACGGGACGACATTTACAGCCTGCAAGTTGT 1680

Qy 1944 TTTTCGGAGTCTGTCCACGGCCCAACGCTCACTTGACTCATGCTTAAGATGCACAA 2003
 Db 1681 TTTTCGGAGTCTGTCCACGGCCCAACGCTCACTTGACTCATGCTTAAGATGCACAA 1740
 Qy 2004 CCGCACCCCATACACAACTACAAGGGCCCAAAAGCAACTGTGGGTAGCTTTCCACAG 2063
 Db 1741 CCGCACCCCATACACAACTACAAGGGCCCAAAAGCAACTGTGGGTAGCTTTCCACAG 1800
 Qy 2064 ACATAAATATAGTCTCATCTGCAATCAAGACATAGCCAGGTAAACCAACTCCCT 2123
 Db 1801 ACATAAATATAGTCTCATCTGCAATCAAGACATAGCCAGGTAAACCAACTCCCT 1860
 Qy 2124 GGATCTGAGCCCAACGCTGGGAGTCTGCTGTCACTTCAAGCCACAGAAACGGCCA 2183
 Db 1861 GGATCTGAGCCCAACGCTGGGAGTCTGCTGTCACTTCAAGCCACAGAAACGGCCA 1920
 Qy 2184 CACATGTTTACAGCTCACACGCCCTCTCCATTTCATCGAACTTCTCAGTGTCCCTGCT 2243
 Db 1921 CACATGTTTACAGCTCACACGCCCTCTCCATTTCATCGAACTTCTCAGTGTCCCTGCT 1980
 Qy 2244 GGTGCTGGCACAGGGAACAGCATGCCCTCCGGGGTTCATGCCACCCAGAGACTGTGC 2303
 Db 1981 GGTGCTGGCACAGGGAACAGCATGCCCTCCGGGGTTCATGCCACCCAGAGACTGTGC 2040
 Qy 2304 TGTCTATGGCCCAACTCATGCTCCCTCTTGGGTACACCACTCTCCAGCTGTGACC 2363
 Db 2041 TGTCTATGGCCCAACTCATGCTCCCTCTTGGGTACACCACTCTCCAGCTGTGACC 2100
 Qy 2364 ACCGATGTCACACACCCCAACCACTTGTCCACACAGTACCCAGTACGACATCGTCC 2423
 Db 2101 ACCGATGTCACACACCCCAACCACTTGTCCACACAGTACCCAGTACGACATCGTCC 2160
 Qy 2424 TGGTCCCGAGATATTTCCACTGACAGCGCCGCCCAACAGAGGACAGTCCCGAG 2483
 Db 2161 TGGTCCCGAGATATTTCCACTGACAGCGCCGCCCAACAGAGGACAGTCCCGAG 2220
 Qy 2484 CCACCTCTGCAACTGACGCCCTCAGTCACTCCCTTTTAAAGCACTGATTCTACCAATG 2543
 Db 2221 CCACCTCTGCAACTGACGCCCTCAGTCACTCCCTTTTAAAGCACTGATTCTACCAATG 2280
 Qy 2544 CAAACACATCTGGGTCTCGATTTATGACAGAGACTTTGGACATACGAGGACCTCAGAC 2603
 Db 2281 CAAACACATCTGGGTCTCGATTTATGACAGAGACTTTGGACATACGAGGACCTCAGAC 2340
 Qy 2604 CGGAGAACACTGCGCCCAACCCCAACAGTGTATGTAAACAGTGGAAAGCGGCCCT 2663
 Db 2341 CGGAGAACACTGCGCCCAACCCCAACAGTGTATGTAAACAGTGGAAAGCGGCCCT 2400
 Qy 2664 GCTGCCCTCCACACACACATACACTCACTGATCTACAGCCCTGTTCCGCGTCAGAG 2723
 Db 2401 GCTGCCCTCCACACACATACACTCACTGATCTACAGCCCTGTTCCGCGTCAGAG 2460
 Qy 2724 TCCCACTAGACCCAGTGGAGGGGTTAGAGCAAGTAGGGGCGAGTTTCCAAATCACC 2783
 Db 2461 TCCCACTAGACCCAGTGGAGGGGTTAGAGCAAGTAGGGGCGAGTTTCCAAATCACC 2520
 Qy 2784 CTGTTCAGGGAGTGGCCGATCTGAGTCTCTTGTGACTTAAGGCTCCGGCTTGGGAAT 2843
 Db 2521 CTGTTCAGGGAGTGGCCGATCTGAGTCTCTTGTGACTTAAGGCTCCGGCTTGGGAAT 2580
 Qy 2844 AAAAGTTTCTTCTGGCCCTTTAGCCCTTAAACCAACCAACCAACCAACCAACCAAC 2886
 Db 2581 AAAAGTTTCTTCTGGCCCTTTAGCCCTTAAACCAACCAACCAACCAACCAACCAAC 2623

RESULT 7
 ACD41826
 ID ACD41826 standard; cDNA; 2623 BP.
 XX
 AC ACD41826;
 XX
 DT 05-SEP-2003 (first entry)

XX DE Human secreted/transmembrane protein (PRO) cDNA #32.
XX KW Human, ss; gene; PRO; secreted protein; transmembrane protein; tumour;
KW proteoglycan; gene therapy; tumour necrosis factor-alpha; TNF-alpha; blood;
KW chondrocyte cell proliferation; FFA; skeletal muscle cell; adipocyte cell;
KW chondrocyte cell proliferation; chondrocyte cell differentiation;
KW pericyte cell; inner ear utricular supporting cell; r-lymphocyte cell;
KW endothelial cell; A-peptide; factor VIIA.
XX OS Homo sapiens.
XX PN US2003036179-A1.
XX PD 20-FEB-2003.
XX PF 10-MAY-2002; 2002US-00142431.
XX PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018624.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 08-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030352.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-0074259.
PR 28-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX (GETH) GENENTECH INC.
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PU, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2003-466355/44.
DR P-PSDB; ABO24872.
XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO1114 or
PT PRO4978, useful in molecular biology, chromosome and gene mapping, in
FT generating antisense RNA and DNA, and in gene therapy.
XX Claim 2; Fig 63; 659pp; English.
XX The invention relates to an isolated nucleic acid comprising at least 80%
CC sequence identity to a PRO (secreted and transmembrane protein) cDNA
CC comprising a nucleic acid (a) encoding a PRO polypeptide, or its
CC extracellular domain (with or without its associated signal peptide),
CC which comprises any of the 275 120-850 residue amino acid sequences,
CC given in the specification; (b) comprising any of the 275 300-3500
CC nucleotide sequences, given in the specification; or (c) comprising the
CC full-length coding sequence of the nucleotide sequences given in the
CC specification, or of the DNA deposited under any of the American Type
CC Culture Collection (ATCC) Accession Numbers listed in the specification.
CC Also included are a vector comprising the novel nucleic acid, a host cell
CC comprising the vector, producing a PRO polypeptide, the isolated PRO


```
Db      1681  TTTCCGGAGTCTGTCACAGCCGACACAGCTTCACTTGACTATGCTGCTAAGATGCACAA 1740
Qy      2004  CCGCACACCCATACACAACTACAAGGGCCACAAGAACTCTGCTGGTTAGCTTTCCACAG 2063
Db      1741  CCGCACACCCATACACAACTACAAGGGCCACAAGAACTCTGCTGGTTAGCTTTCCACAG 1800
Qy      2064  ACATAATATAGTCCATCTGCAATCAAGACATAGCCAGGTAAACCAACTCCCT 2123
Db      1801  ACATAATATAGTCCATCTGCAATCAAGACATAGCCAGGTAAACCAACTCCCT 1860
Qy      2124  GGATCTGCAGCCACACAGTGGAGTCTGGCTGTCACTTCAAGCCACAGAAACGGCCA 2183
Db      1861  GGATCTGCAGCCACACAGTGGAGTCTGGCTGTCACTTCAAGCCACAGAAACGGCCA 1920
Qy      2184  CACATGTTACAGCTCACAGCCCTCTCCATTCATTCGAATTCAGTGTCCCTGTCCCT 2243
Db      1921  CACATGTTACAGCTCACAGCCCTCTCCATTCATTCGAATTCAGTGTCCCTGTCCCT 1980
Qy      2244  GGTGCTGGCACAGGACAGCAGTCCCTCCGCGGTATGCCACCCAGAGACTGTGCG 2303
Db      1981  GGTGCTGGCACAGGACAGCAGTCCCTCCGCGGTATGCCACCCAGAGACTGTGCG 2040
Qy      2304  TGTCTATGGCCCCCACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC 2363
Db      2041  TGTCTATGGCCCCCACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC 2100
Qy      2364  ACCGATGTCACACACCCCAACCACTTGTGTCACACAGCTACCCAGTACGACATCGTCC 2423
Db      2101  ACCGATGTCACACACCCCAACCACTTGTGTCACACAGCTACCCAGTACGACATCGTCC 2160
Qy      2424  TGGCTCCCGAGAGTATCTCCCACTGAGACAGCGCGCCACAGAGGCAAGTCCCGAC 2483
Db      2161  TGGCTCCCGAGAGTATCTCCCACTGAGACAGCGCGCCACAGAGGCAAGTCCCGAC 2220
Qy      2484  CCACCTCTGCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCTTGAATTCACAAATG 2543
Db      2221  CCACCTCTGCAACTGCAGCCCTCAGTCAACCCCTTTTAAAGCACCTTGAATTCACAAATG 2280
Qy      2544  CAAACACATCTGGTCTGGATTATGACAGAGACTTTGGACATACAGGACCCCTCAGAC 2603
Db      2281  CAAACACATCTGGTCTGGATTATGACAGAGACTTTGGACATACAGGACCCCTCAGAC 2340
Qy      2604  CGGAGGAACACTCCCAACCCCAACAGCTGCTTATGTAACACAGTGAAGCGGCCCT 2663
Db      2341  CGGAGGAACACTCCCAACCCCAACAGCTGCTTATGTAACAGTGAAGCGGCCCT 2400
Qy      2664  GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTTCGGCGTCAAG 2723
Db      2401  GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTTCGGCGTCAAG 2460
Qy      2724  TCCCACTAGACCCAGTGGAGGGTTAGACACCAAGTAGGGGCGAGTTTCCAATTCAAC 2783
Db      2461  TCCCACTAGACCCAGTGGAGGGTTAGACACCAAGTAGGGGCGAGTTTCCAATTCAAC 2520
Qy      2784  CTGTCAGGAGTGAAGCGGATCTCAGCTTCTGCTGACTTAAGGTCGGCTTGGGAATT 2843
Db      2521  CTGTCAGGAGTGAAGCGGATCTCAGCTTCTGCTGACTTAAGGTCGGCTTGGGAATT 2580
Qy      2844  AAAGTTTGTTCGGCTTTAGCCATAAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGG 2886
Db      2581  AAAGTTTGTTCGGCTTTAGCCATAAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGG 2623
```

RESULT 8

ACA04055

ID ACA04055 standard; cDNA; 2623 BP.

XX

AC ACA04055;

XX

DT 27-MAY-2003 (first entry)

XX

DE Human cDNA encoding a secreted/transmembrane protein, SEQ ID 63.

XX Human; ss; gene; secreted protein; transmembrane protein; PRO; inflammatory disease; organ failure; atherosclerosis; cardiac injury; infertility; birth defects; premature aging; AIDS; biosensor; acquired immunodeficiency syndrome; cancer; diabetic complication; bioreactor; tumour.

Homo sapiens.

US2003032155-A1.

13-FEB-2003.

03-MAY-2002; 2002US-00137865.

31-MAR-1997; 97WO-US005230.

12-JUN-1998; 98WO-US012456.

14-JUL-1998; 98WO-US014552.

28-AUG-1998; 98WO-US017888.

10-SEP-1998; 98WO-US018824.

14-SEP-1998; 98WO-US019093.

14-SEP-1998; 98WO-US019094.

16-SEP-1998; 98WO-US019177.

17-SEP-1998; 98WO-US019330.

07-OCT-1998; 98WO-US019437.

29-OCT-1998; 98WO-US021141.

29-OCT-1998; 98WO-US022991.

20-NOV-1998; 98WO-US024855.

01-DEC-1998; 98WO-US025108.

05-JAN-1999; 99WO-US000106.

08-MAR-1999; 99WO-US005028.

10-MAR-1999; 99WO-US005190.

20-APR-1999; 99WO-US008615.

14-MAY-1999; 99WO-US010733.

02-JUN-1999; 99WO-US012252.

01-SEP-1999; 99WO-US020111.

08-SEP-1999; 99WO-US020594.

13-SEP-1999; 99WO-US020944.

15-SEP-1999; 99WO-US021090.

05-OCT-1999; 99WO-US021547.

29-NOV-1999; 99WO-US023089.

30-NOV-1999; 99WO-US028214.

30-NOV-1999; 99WO-US028313.

01-DEC-1999; 99WO-US028409.

01-DEC-1999; 99WO-US028301.

02-DEC-1999; 99WO-US028634.

02-DEC-1999; 99WO-US028551.

02-DEC-1999; 99WO-US028564.

16-DEC-1999; 99WO-US028565.

20-DEC-1999; 99WO-US030095.

20-DEC-1999; 99WO-US030911.

22-DEC-1999; 99WO-US030999.

30-DEC-1999; 99WO-US030720.

30-DEC-1999; 99WO-US031243.

05-JAN-2000; 99WO-US031274.

06-JAN-2000; 2000WO-US000219.

06-JAN-2000; 2000WO-US000277.

11-FEB-2000; 2000WO-US000376.

18-FEB-2000; 2000WO-US003565.

18-FEB-2000; 2000WO-US004341.

22-FEB-2000; 2000WO-US004414.

24-FEB-2000; 2000WO-US004914.

24-FEB-2000; 2000WO-US005004.

01-MAR-2000; 2000WO-US005601.

02-MAR-2000; 2000WO-US005746.

02-MAR-2000; 2000WO-US005841.

10-MAR-2000; 2000WO-US006319.

15-MAR-2000; 2000WO-US006884.

20-MAR-2000; 2000WO-US007377.

21-MAR-2000; 2000WO-US007532.

30-MAR-2000; 2000WO-US008439.

PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023525.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2000US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00815744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.

(GETH) GENENTECH INC.

XX Baker KP, Bersesini M, Deforge L, Desnoyers L, Filvaroff E, Cao W;
PI Gerritsen MB, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX

DR WPI; 2003-331925/31.

DR P-PSDB; ABU66877.

XX New secreted and transmembrane nucleic acids and polypeptides, designated
PT as PRO, useful for treating inflammation, organ failure, atherosclerosis,
PT cardiac injury, infertility, birth defects, premature aging, AIDS, or
PT cancer.

PS Claim 2; Fig 63; 659pp; English.

XX The invention relates to an isolated nucleic acid comprising, or which is
CC at least 80% identical to, or the full-length coding sequence of, any of
CC the 275 nucleotide sequences, encoding the corresponding PRO polypeptide
CC (one of 275 secreted or transmembrane proteins). The nucleic acid further
CC comprises the full-length coding sequence of the DNA deposited under
CC American Type Culture Collection (ATCC) accession number in a list given
CC in the specification. Also included are vectors and host cells for
CC producing PRO proteins, PRO fusion proteins, anti-PRO antibodies, PRO
CC extracellular domains and mature sequences, methods of detecting PRO
CC proteins, methods for stimulating the release of TNF-alpha (tumour
CC necrosis factor alpha) from human blood, (and the proliferation of
CC differentiation of chondrocyte cells, the proliferation of, or gene
CC expression in pericyte cells, the release of proteoglycans from
CC cartilage, proliferation of inner ear articular supporting cells, the
CC proliferation of T-lymphocyte cells, the release of a cytokine from

CC peripheral blood mononuclear cells (PBMC), or the proliferation of
CC endothelial cells), a method for modulating the uptake of glucose or free
CC fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the
CC binding of A-peptide to factor VIIA, or the differentiation of adipocyte
CC cells, a method for detecting the presence of a tumour in a mammal and an
CC oligonucleotide probe derived from any of the nucleotide sequences cited
CC above. The nucleic acids and polypeptides are useful for treating
CC inflammatory diseases, organ failure, atherosclerosis, cardiac injury,
CC infertility, birth defects, premature aging, AIDS (acquired
CC immunodeficiency syndrome), cancer, or diabetic complications. The
CC nucleic acids are useful as hybridisation probes, in chromosome and gene
CC mapping, and in generating antisense RNA or DNA. The polypeptides are
CC useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both
CC are useful in tissue typing. The present sequence encodes a PRO protein
CC of the invention

XX Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Query Match 89.0%; Score 2621.4; DB 7; Length 2623;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 264 GCGGAGCGCTGGGAGAGAGAGAGCGACCTGCCGAGATGGAGCGACCGCACCTG 323
Db 1 GCGGAGCGCTGGGAGAGAGAGAGCGACCTGCCGAGATGGAGCGACCGCACCTG 60
QY 324 GCGGCTGCTGGGCGCTGGCGCTGCTCTGCTGTGACGCTGGCGCTGTCGGGACCA 383
Db 61 GCGGCTGCTGGCGCTGGCGCTGCTCTGCTGTGACGCTGGCGCTGTCGGGACCA 120
QY 384 GCGCGGAGGCGACCTGCCCGCGCGCGCGCGCTACCCACTGCTGGGAAACCTCTCGCA 443
Db 121 GCGCGGAGGCGACCTGCCCGCGCGCGCGCGCTACCCACTGCTGGGAAACCTCTCGCA 180
QY 444 GCTACGCGCGCGCGCGCTGTATTTCAGGGCTCATTCGCGCTGAGTAAGATACGACCGGT 503
Db 181 GCTACGCGCGCGCGCGCTGTATTTCAGGGCTCATTCGCGCTGAGTAAGATACGACCGGT 240
QY 504 GTTCACCATCTACCTGGGACCTGGCGCGCTGTGGTGGTCTGTTGGGCGAGAGCTGT 563
Db 241 GTTCACCATCTACCTGGGACCTGGCGCGCTGTGGTGGTCTGTTGGGCGAGAGCTGT 300
QY 564 GCGGAGGCGCTGGGAGGTGAGGTGAGGTTCAGCGCGCGGGGAAACCTGATGCT 623
Db 301 GCGGAGGCGCTGGGAGGTGAGGTGAGGTTCAGCGCGCGGGGAAACCTGATGCT 360
QY 624 GGAAGGACTTTTGATGGCAATGGGGTTTCTTCTCCAAACGGGAGCGGTGGAGGAGCT 683
Db 361 GGAAGGACTTTTGATGGCAATGGGGTTTCTTCTCCAAACGGGAGCGGTGGAGGAGCT 420
QY 684 GAGGAAGTTTACCATGCTTCTGCGGACCTGGGATGGGATGGGAGAGGCGAGGA 743
Db 421 GAGGAAGTTTACCATGCTTCTGCGGACCTGGGATGGGATGGGAGAGGCGAGGA 480
QY 744 GCTGATCCAGGGGAGGCGCGGTGTCTGCTGGAGACATTCCAGGGGACAGAGAGCGCCC 803
Db 481 GCTGATCCAGGGGAGGCGCGGTGTCTGCTGGAGACATTCCAGGGGACAGAGAGCGCCC 540
QY 804 ATTGATTCCTCTGCTGCTGGCGCGCGACCTCCAAAGCTAGTGTCTCTCTCTCTT 863
Db 541 ATTGATTCCTCTGCTGCTGGCGCGCGACCTCCAAAGCTAGTGTCTCTCTCTCTT 600
QY 864 TGSCCTCCGCTTCTCTATGAGGATAAGGAGTTCCAGGCGGTGTCGGGCGAGCTGGTG 923
Db 601 TGSCCTCCGCTTCTCTATGAGGATAAGGAGTTCCAGGCGGTGTCGGGCGAGCTGGTG 660
QY 924 TACCCTGTGGGAGTCACTCCAGGGGGTCCAGACTACGAGATTTCTCTGGTTCCT 983
Db 661 TACCCTGTGGGAGTCACTCCAGGGGGTCCAGACTACGAGATTTCTCTGGTTCCT 720
QY 984 GCGGCGCGCTGCCAGGCGCGCGCGCGCGCTCTCCACAGCTCAGCACCTTGTGCTGCTT 1043
Db 721 GCGGCGCGCTGCCAGGCGCGCGCGCGCGCTCTCCACAGCTCAGCACCTTGTGCTGCTT 780

QY 1044 CACAGTCCGGCAGGTGACAGACACACAGGGGAACTGTGATGCTTCGGGCCCCGACAGTGA 1103
Db 781 CACAGTCCGGCAGGTGACAGACACACAGGGGAACTGTGATGCTTCGGGCCCCGACAGTGA 840
QY 1104 CTTTGTCCGATCCCTTCCTGCTGAAGATGGCAAGAGGAAACAAACCCAGGACACAGAAAT 1163
Db 841 CTTTGTCCGATCCCTTCCTGCTGAAGATGGCAAGAGGAAACAAACCCAGGACACAGAAAT 900
QY 1164 CACCAACAAGAAATGCTGATGACAGTCAATTAATTTGCTGTTGCTGGGACGATGACGGT 1223
Db 901 CACCAACAAGAAATGCTGATGACAGTCAATTAATTTGCTGTTGCTGGGACGATGACGGT 960
QY 1224 CAGCACACAGGTCCGGCTATACCTCCCTGCTCCTGATGAAATACACCTCATGTCCAAAGTG 1283
Db 961 CAGCACACAGGTCCGGCTATACCTCCCTGCTCCTGATGAAATACACCTCATGTCCAAAGTG 1020
QY 1284 GGTACGTGAGGAGCTGAATCCGGAGCTGGGGGCTGGCCAGGACCAACAGCCTTAGGGGACCG 1343
Db 1021 GGTACGTGAGGAGCTGAATCCGGAGCTGGGGGCTGGCCAGGACCAACAGCCTTAGGGGACCG 1080
QY 1344 TACCCGCTCCCTTACACCGACCGGTTCTGCATGAGCGCAGCGGCTGTGGCGCTGGT 1403
Db 1081 TACCCGCTCCCTTACACCGACCGGTTCTGCATGAGCGCGAGCGGCTGTGGCGCTGGT 1140
QY 1404 GCCCATGGGAATACCCCGCACCTCATGCGGACCAACCGCTTCGAGGGTACACCTGCC 1463
Db 1141 GCCCATGGGAATACCCCGCACCTCATGCGGACCAACCGCTTCGAGGGTACACCTGCC 1200
QY 1464 CCAGGGCAACGAGGTCTTCCCTCCCTTGGCTCCATCTGCATGACCCCAACATCTTCAA 1523
Db 1201 CCAGGGCAACGAGGTCTTCCCTCCCTTGGCTCCATCTGCATGACCCCAACATCTTCAA 1260
QY 1524 GCACCCAGAAGGTTAAACCCAGACCGTTTCCTGGATGCAGATGGAACGTTTCAGGAAGCA 1583
Db 1261 GCACCCAGAAGGTTAAACCCAGACCGTTTCCTGGATGCAGATGGAACGTTTCAGGAAGCA 1320
QY 1584 TGAGGCGTTCCTGCCTCTCTTAGGAGACGCTGTGCTGCTTGGAGAGGCGCTGGCAA 1643
Db 1321 TGAGGCGTTCCTGCCTCTCTTAGGAGACGCTGTGCTGCTTGGAGAGGCGCTGGCAA 1380
QY 1644 AGCGAGCTCTTCTCTTTCACCAACCTCTCAAGACCTTCTCCCTGGAGAGCGCGTG 1703
Db 1381 AGCGAGCTCTTCTCTTTCACCAACCTCTCAAGACCTTCTCCCTGGAGAGCGCGTG 1440
QY 1704 CCGCCGGACACCTGAGCTCAAGCCACCGTCAAGTGGGCTTTTCAACATTCGCCCAGC 1763
Db 1441 CCGCCGGACACCTGAGCTCAAGCCACCGTCAAGTGGGCTTTTCAACATTCGCCCAGC 1500
QY 1764 CTTCCAGCTGCAAGTCCGTCCTCACTGACCTTCACTCCACACGACGACACAGTGAAGGAA 1823
Db 1501 CTTCCAGCTGCAAGTCCGTCCTCACTGACCTTCACTCCACACGACGACACAGTGAAGGAA 1560
QY 1824 GGCAACTTGAAGTGGTGGTGCCAGGACGCTGCTCCAGCCTCAACAGTGGGCATGGA 1883
Db 1561 GGCAACTTGAAGTGGTGGTGCCAGGACGCTGCTCCAGCCTCAACAGTGGGCATGGA 1620
QY 1884 CAGGGTTAATGCTCTCAGAGTGATACCTGACGAGGACCAATTTACACGCTGCAAGTTGT 1943
Db 1621 CAGGGTTAATGCTCTCAGAGTGATACCTGACGAGGACCAATTTACACGCTGCAAGTTGT 1680
QY 1944 TTTCCGAGTCTGTCCACGCGCCACACGCTCACTTGACTCATGCTGCTAAGATGACAA 2003
Db 1681 TTTCCGAGTCTGTCCACGCGCCACACGCTCACTTGACTCATGCTGCTAAGATGACAA 1740
QY 2004 CCGCACACCCATACAACTACAAGGGCCCAAGCAACTGCTGGGTAGCTTTCACAG 2063
Db 1741 CCGCACACCCATACAACTACAAGGGCCCAAGCAACTGCTGGGTAGCTTTCACAG 1800
QY 2064 ACATAATATAGTCCATCTGCAATCAAGACATAGCCAGGTAAACCACTCCCT 2123
Db 1801 ACATAATATAGTCCATCTGCAATCAAGACATAGCCAGGTAAACCACTCCCT 1860

QY 2124 GGATCTGACGCCACACAGTGGGAGTCTGGCTGTCACTTCAAGGCAACAGAAAGCGCA 2183
Db 1861 GGATCTGACGCCACACAGTGGGAGTCTGGCTGTCACTTCAAGGCAACAGAAAGCGCA 1920
QY 2184 CACATGTTACAGCTCAACAGCCCTCTCCATTTCATCGAACTTCTCAGTGTCCCTGCTCC 2243
Db 1921 CACATGTTACAGCTCAACAGCCCTCTCCATTTCATCGAACTTCTCAGTGTCCCTGCTCC 1980
QY 2244 GGTCCCTGGCACAGGGAACAGCATGCCCCCTTCGCGGTTCATGCCACCCAGAGACTGTGC 2303
Db 1981 GGTCCCTGGCACAGGGAACAGCATGCCCCCTTCGCGGTTCATGCCACCCAGAGACTGTGC 2040
QY 2304 TGTCTATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGACC 2363
Db 2041 TGTCTATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGACC 2100
QY 2364 ACCGATGTCACACACACCCCAACACTTGTCCACACAGCTACCCAGCTACGACATCGTCC 2423
Db 2101 ACCGATGTCACACACACCCCAACACTTGTCCACACAGCTACCCAGCTACGACATCGTCC 2160
QY 2424 TGGCTCCCGAGAGTATCTTCCACTGAGACACGCGCCCCCACAGAGGACAGTCCCCAG 2483
Db 2161 TGGCTCCCGAGAGTATCTTCCACTGAGACACGCGCCCCCACAGAGGACAGTCCCCAG 2220
QY 2484 CCACCTCTGCAACTGACGCGCTCAGTCAACCTTTTAAAGCACCTGATTTACAAATG 2543
Db 2221 CCACCTCTGCAACTGACGCGCTCAGTCAACCTTTTAAAGCACCTGATTTACAAATG 2280
QY 2544 CAAACACATCTGGGCTCTGCGATTATGACACAGAGACTTTGGACATACGAGGACCTCAGAC 2603
Db 2281 CAAACACATCTGGGCTCTGCGATTATGACACAGAGACTTTGGACATACGAGGACCTCAGAC 2340
QY 2604 CGAGGAACACCTGCCCCAACCCCAACACGCTTATGTAAACACGCTGTAAGCGGCCCT 2663
Db 2341 CGAGGAACACCTGCCCCAACCCCAACACGCTTATGTAAACACGCTGTAAGCGGCCCT 2400
QY 2664 GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGGGCGTACAG 2723
Db 2401 GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTGGGCGTACAG 2460
QY 2724 TCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTACGAGGCGAGTTCCAAATTCACC 2783
Db 2461 TCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTACGAGGCGAGTTCCAAATTCACC 2520
QY 2784 CTGTGAGGAGTACGCGATCTGACGTTCTTGTGACTTAAAGGTCGCGCTTGGGAAT 2843
Db 2521 CTGTGAGGAGTACGCGATCTGACGTTCTTGTGACTTAAAGGTCGCGCTTGGGAAT 2580
QY 2844 AAAGTTGTTTCTGCGCTTTAGCCTAAAAAAGGAAAAA 2886
Db 2581 AAAGTTGTTTCTGCGCTTTAGCCTAAAAAAGGAAAAA 2623

RESULT 9
ADA45582
ID ADA45582 standard; cDNA; 2623 BP.
XX
AC ADA45582;
XX
DT 20-NOV-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1906 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; gene; ss;
KW Tumour necrosis factor alpha release; TNF-alpha release;
KW Glucose uptake modulator; FFA uptake modulator;
KW cell proliferation stimulator; cell differentiation stimulator;
KW cell differentiation inhibitor; cytokine release stimulator; tumour;
KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; chromosome mapping; gene mapping;
KW gene therapy; chromosome identification; chromosome marker.
OS Homo sapiens.

XX US2003022328-A1.
 PN
 XX 30-JAN-2003.
 PD
 XX 16-APR-2002; 2002US-00123904.
 XX
 XX 31-MAR-1997; 97WO-US005230.
 PR 12-JUN-1998; 98WO-US012456.
 PR 14-JUL-1998; 98WO-US014552.
 PR 28-AUG-1998; 98WO-US017888.
 PR 10-SEP-1998; 98WO-US018824.
 PR 14-SEP-1998; 98WO-US019093.
 PR 14-SEP-1998; 98WO-US019094.
 PR 14-SEP-1998; 98WO-US019177.
 PR 16-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98WO-US019437.
 PR 07-OCT-1998; 98WO-US021141.
 PR 29-OCT-1998; 98WO-US022991.
 PR 29-OCT-1998; 98WO-US022992.
 PR 20-NOV-1998; 98WO-US024855.
 PR 01-DEC-1998; 98WO-US025108.
 PR 05-JAN-1999; 99WO-US000106.
 PR 08-MAR-1999; 99WO-US005028.
 PR 10-MAR-1999; 99WO-US005190.
 PR 20-APR-1999; 99WO-US008615.
 PR 14-MAY-1999; 99WO-US010733.
 PR 02-JUN-1999; 99WO-US012252.
 PR 01-SEP-1999; 99WO-US020111.
 PR 08-SEP-1999; 99WO-US020594.
 PR 13-SEP-1999; 99WO-US020944.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.
 PR 05-OCT-1999; 99WO-US023089.
 PR 29-NOV-1999; 99WO-US028214.
 PR 30-NOV-1999; 99WO-US028313.
 PR 30-NOV-1999; 99WO-US028409.
 PR 01-DEC-1999; 99WO-US028301.
 PR 01-DEC-1999; 99WO-US028634.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028564.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 20-DEC-1999; 99WO-US030999.
 PR 22-DEC-1999; 99WO-US030720.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005746.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 20-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000US-00747259.
 PR 20-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001US-00796498.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 01-MAR-2001; 2001WO-US006666.
 PR 09-MAR-2001; 2001US-00802706.
 PR 14-MAR-2001; 2001US-00806689.
 PR 22-MAR-2001; 2001US-00816744.
 PR 05-APR-2001; 2001US-00828366.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854280.
 PR 18-MAY-2001; 2001US-00860216.
 PR 25-MAY-2001; 2001US-00866028.
 PR 25-MAY-2001; 2001US-00866034.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 21-JUN-2001; 2001US-00887879.
 PR 22-JUN-2001; 2001WO-US020116.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 18-JUL-2001; 2001US-00908827.
 PR 06-AUG-2001; 2001US-00924419.
 PR 09-AUG-2001; 2001US-00927796.
 PR 16-AUG-2001; 2001US-00931836.
 PR 19-DEC-2001; 2001US-00028072.
 XX
 XX (GETH) GENENTECH INC.
 PA
 XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 XX WPI; 2003-584997/55.
 DR P-PSDB; ADA45583.
 DR
 XX Novel secreted and transmembrane polypeptide for modulating biological
 PT activity of cell expressing the polypeptide, identifying agonists or
 PT antagonists of polypeptide, and as molecular weight markers.
 XX
 XX Claim 2; Fig 63; 659pp; English.
 XX
 CC The invention describes 305 nucleic acids encoding PRO (secreted and
 CC transmembrane) polypeptides (I). (I) is useful for stimulating the
 CC release of TNF-alpha from human blood, for modulating the uptake of
 CC glucose or FFA by skeletal muscle cells or adipocyte cells, for
 CC stimulating the proliferation or differentiation of chondrocyte cells,
 CC for stimulating the proliferation of or gene expression in pericyte
 CC cells, for stimulating the release of proteoglycans from cartilage, for
 CC stimulating the proliferation of inner ear utricular supporting cells,
 CC for stimulating the proliferation of T-lymphocyte cells, for stimulating
 CC the release of a cytokine from PBMC cells, for inhibiting the binding of
 CC A-peptide to factor VIIa, for inhibiting the differentiation of adipocyte
 CC cells, for stimulating proliferation of endothelial cells, for detecting
 CC the presence of tumour in a mammal. The tumour is lung, colon, breast,
 CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes
 CC are useful for isolating genomic and cDNA nucleotide sequences or
 CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful
 CC in assays to identify other proteins or molecules involved in binding
 CC interaction. A polynucleotide (ii) encoding (I) is useful in chromosome
 CC and gene mapping, in generation of antisense RNA and DNA, in the
 CC preparation of PRO polypeptide, for generating transgenic animals or
 CC knockout animals which in turn are useful in the development and
 CC screening of therapeutically useful reagents, in gene therapy, for
 CC chromosome identification, as chromosome marker, and for generating
 CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.

QY	2304	TGTCATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC	2363	PR	14-JUL-1998;	98WO-US014552.
PR				PR	28-AUG-1998;	98WO-US017898.
Db	2041	TGTCATGGCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCAGCCTGTGACC	2100	PR	10-SEP-1998;	98WO-US018824.
PR				PR	14-SEP-1998;	98WO-US019093.
QY	2364	ACCGATGTCCACACACACCCCAACCACTTGTCCACACAGCTACCCACGCTACGACATCGTGCC	2423	PR	14-SEP-1998;	98WO-US019177.
PR				PR	16-SEP-1998;	98WO-US019330.
Db	2101	ACCGATGTCCACACACACCCCAACCACTTGTCCACACAGCTACCCACGCTACACATCGTGCC	2160	PR	17-SEP-1998;	98WO-US019437.
PR				PR	07-OCT-1998;	98WO-US021141.
QY	2424	TGGCTCCCGAGATATCTTCCCACTGAGACACGCGCGCCCCACAGGACACAGTCCCGCAG	2483	PR	29-OCT-1998;	98WO-US022991.
PR				PR	29-OCT-1998;	98WO-US022992.
Db	2161	TGGCTCCCGAGATATCTTCCCACTGAGACACGCGCGCCCCACAGGACACAGTCCCGCAG	2220	PR	20-NOV-1998;	98WO-US024855.
PR				PR	01-DEC-1998;	98WO-US025108.
QY	2484	CCACCTCTGCACTGAGCGCCCTCAGTCACCCCTTTTAAAGCACCCCTGATTCTACCAATG	2543	PR	05-JAN-1999;	99WO-US000106.
PR				PR	08-MAR-1999;	99WO-US005028.
Db	2221	CCACCTCTGCACTGAGCGCCCTCAGTCACCCCTTTTAAAGCACCCCTGATTCTACCAATG	2280	PR	10-MAR-1999;	99WO-US005190.
PR				PR	20-APR-1999;	99WO-US008615.
QY	2544	CAACACATCTGGGCTCGGATTATGACACAGAGACTTTTGGACATACGAGGACCCCTCAGAC	2603	PR	14-MAY-1999;	99WO-US010733.
PR				PR	02-JUN-1999;	99WO-US012252.
Db	2281	CAACACATCTGGGCTCGGATTATGACACAGAGACTTTTGGACATACGAGGACCCCTCAGAC	2340	PR	01-SEP-1999;	99WO-US020111.
PR				PR	08-SEP-1999;	99WO-US020594.
QY	2604	CGGAGAACACCTGCGCCCAACCCCAACACGTCGTTATGTAACCACTGTAAGCGGCGCCCT	2663	PR	13-SEP-1999;	99WO-US020944.
PR				PR	15-SEP-1999;	99WO-US021090.
Db	2341	CGGAGAACACCTGCGCCCAACCCCAACACGTCGTTATGTAACCACTGTAAGCGGCGCCCT	2400	PR	15-SEP-1999;	99WO-US021547.
PR				PR	05-OCT-1999;	99WO-US023089.
QY	2664	GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTTCGGCGTCAGAG	2723	PR	29-NOV-1999;	99WO-US028214.
PR				PR	30-NOV-1999;	99WO-US028313.
Db	2401	GCTGCCCTCCACACACATACACACTCACTGATCTACAGCCCTGTTTCGGCGTCAGAG	2460	PR	30-NOV-1999;	99WO-US028409.
PR				PR	01-DEC-1999;	99WO-US028301.
QY	2724	TCCCACTAGACCCAGTGAAGGGTTAGAGCAAGTAGGGGCCCAATTTCCAATTCACC	2783	PR	01-DEC-1999;	99WO-US028634.
PR				PR	02-DEC-1999;	99WO-US028551.
Db	2461	TCCCACTAGACCCAGTGAAGGGTTAGAGCAAGTAGGGGCCCAATTTCCAATTCACC	2520	PR	02-DEC-1999;	99WO-US028564.
PR				PR	02-DEC-1999;	99WO-US028565.
QY	2784	CTGTAGGGAGTAGCCGAGCTGACGTTCTTGTGACTTAAGGTCGCGCTTGGGAATT	2843	PR	16-DEC-1999;	99WO-US030095.
PR				PR	20-DEC-1999;	99WO-US030911.
Db	2521	CTGTAGGGAGTAGCCGAGCTGACGTTCTTGTGACTTAAGGTCGCGCTTGGGAATT	2580	PR	20-DEC-1999;	99WO-US030999.
PR				PR	22-DEC-1999;	99WO-US030720.
QY	2844	AAAGTTTGTCTTGGCCTTTAGCCTAAAAAATAAAAAAAAAA	2886	PR	30-DEC-1999;	99WO-US031243.
PR				PR	05-JAN-2000;	2000WO-US000219.
Db	2581	AAAGTTTGTCTTGGCCTTTAGCCTAAAAAATAAAAAAAAAA	2623	PR	06-JAN-2000;	2000WO-US000277.
PR				PR	06-JAN-2000;	2000WO-US000376.
RESULT 10				PR	11-FEB-2000;	2000WO-US003565.
ADA76013				PR	18-FEB-2000;	2000WO-US004341.
ID ADA76013 standard; cDNA; 2623 BP.				PR	18-FEB-2000;	2000WO-US004342.
XX				PR	22-FEB-2000;	2000WO-US004414.
AC ADA76013;				PR	24-FEB-2000;	2000WO-US004914.
XX				PR	24-FEB-2000;	2000WO-US005004.
XX				PR	01-MAR-2000;	2000WO-US005601.
XX				PR	02-MAR-2000;	2000WO-US005746.
DE				PR	02-MAR-2000;	2000WO-US005841.
XX				PR	10-MAR-2000;	2000WO-US006319.
XX				PR	15-MAR-2000;	2000WO-US006884.
KW				PR	20-MAR-2000;	2000WO-US007377.
KW				PR	21-MAR-2000;	2000WO-US007532.
KW				PR	30-MAR-2000;	2000WO-US008439.
KW				PR	17-MAY-2000;	2000WO-US013705.
KW				PR	22-MAY-2000;	2000WO-US014042.
KW				PR	30-MAY-2000;	2000WO-US014941.
KW				PR	02-JUN-2000;	2000WO-US015264.
KW				PR	28-JUL-2000;	2000WO-US020710.
KW				PR	11-AUG-2000;	2000WO-US022031.
KW				PR	23-AUG-2000;	2000WO-US023522.
XX				PR	24-AUG-2000;	2000WO-US023328.
OS				PR	08-NOV-2000;	2000WO-US030952.
XX				PR	10-NOV-2000;	2000WO-US030873.
PN				PR	01-DEC-2000;	2000WO-US032678.
XX				PR	20-DEC-2000;	2000US-00747259.
PD				PR	20-DEC-2000;	2000WO-US034956.
XX				PR	28-FEB-2001;	2001US-00796498.
PF				PR	28-FEB-2001;	2001WO-US006520.
XX				PR	01-MAR-2001;	2001WO-US006666.
PR				PR	09-MAR-2001;	2001US-00802706.

Db 961 CAGCACACGGTGGGCTATACCTCTCTCTGATGAATACCTCTATGTCCTCAAAAGTG 1020
Qy 1284 GGTAGTGAAGCTGAATCGGAGCTGGGGCTGGCCAGGACCAAGCCTAGGGAGCG 1343
Db 1021 GGTAGTGAAGCTGAATCGGAGCTGGGGCTGGCCAGGACCAAGCCTAGGGAGCG 1080
Qy 1344 TACCGCTCTCCCTTACACCGACGGTCTGTGATGAGCGGAGCGGCTGCTGGCGTGGT 1403
Db 1081 TACCGCTCTCCCTTACACCGACGGTCTGTGATGAGCGGAGCGGCTGCTGGCGTGGT 1140
Qy 1404 GCCATGGGAATPACCCGACCCCTCATGCGGACCAACCGCTTCCAGGGTACACCTGCC 1523
Db 1141 GCCATGGGAATPACCCGACCCCTCATGCGGACCAACCGCTTCCAGGGTACACCTGCC 1200
Qy 1464 CCAGGCGAGGAGTCTTCCCTCTTGGCTTCATCTGATGACCCGACCCCAATCTTCAA 1523
Db 1201 CCAGGCGAGGAGTCTTCCCTCTTGGCTTCATCTGATGACCCGACCCCAATCTTCAA 1260
Qy 1524 GCACCCAGAAGATTCACCCAGACCGTTTCTGATGACGATGGATGGGATTCAGGAAGCA 1583
Db 1261 GCACCCAGAAGATTCACCCAGACCGTTTCTGATGACGATGGGATTCAGGAAGCA 1320
Qy 1584 TGAGCGTTCGCTCTCTCTTCTTAGGGAAGCGTGTGCTGCTTGGAGGGCTGGCAAA 1643
Db 1321 TGAGCGTTCGCTCTCTCTTCTTAGGGAAGCGTGTGCTGCTTGGAGGGCTGGCAAA 1380
Qy 1644 AGCGAGCTCTCTCTCTTCCACACCATCTTCAAGGCTTCTCCCTGGAGAGCCGCTG 1703
Db 1381 AGCGAGCTCTCTCTCTTCTTCCACACCATCTTCAAGGCTTCTCCCTGGAGAGCCGCTG 1440
Qy 1704 CCCGCGGACACCTCGAGCTCAAGCCACCGTCTAGTGGCTTTTCAACATTCCTCCAGC 1763
Db 1441 CCCGCGGACACCTCGAGCTCAAGCCACCGTCTAGTGGCTTTTCAACATTCCTCCAGC 1500
Qy 1764 CTTCAGCTGGAAGTCCGCTCCATGACCTTCAACACGACGATGAGTGAAGAA 1823
Db 1501 CTTCAGCTGGAAGTCCGCTCCATGACCTTCAACACGACGATGAGTGAAGAA 1560
Qy 1824 GGCMACTTGAAGTGGTGGTCCAGGAGCGTGGCTCCAGCTCAACAGTGGGATGGA 1883
Db 1561 GGCMACTTGAAGTGGTGGTCCAGGAGCGTGGCTCCAGCTCAACAGTGGGATGGA 1620
Qy 1884 CAGGGTTAATGTCTCCAGAGTGTACACTGCAGGACGCCACATTTACACGCTCGAGTTGT 1943
Db 1621 CAGGGTTAATGTCTCCAGAGTGTACACTGCAGGAGGCCACATTTACACGCTCGAGTTGT 1680
Qy 1944 TTTCCGGAGTGTGCCAGGCCACACGCTCATCTGATGATGCTTAAGATGACAAA 2003
Db 1681 TTTCCGGAGTGTGCCAGGCCACACGCTCATCTGATGATGCTTAAGATGACAAA 1740
Qy 2004 CCGCACACCCATACAACTACAAGGGCCACAAAGCACTGTGGTTAGCTTTCCACAG 2063
Db 1741 CCGCACACCCATACAACTACAAGGGCCACAAAGCACTGTGGTTAGCTTTCCACAG 1800
Qy 2064 ACATAAATAGTTCATCTGCAATCACAAGACCATAGCCAGGTAAACCAACTCCCT 2123
Db 1801 ACATAAATAGTTCATCTGCAATCACAAGACCATAGCCAGGTAAACCAACTCCCT 1860
Qy 2124 GGATCTGCAGCCACACGCTGGAGTGTGGCTGACCTTCAAGCCACAGAAAAGGCCA 2183
Db 1861 GGATCTGCAGCCACACGCTGGAGTGTGGCTGACCTTCAAGCCACAGAAAAGGCCA 1920
Qy 2184 CACATGTTTACAGCTCACACGCTCTCCATTCATCGAACTTCTCAGTGTCTGCTGCC 2243
Db 1921 CACATGTTTACAGCTCACACGCTCTCCATTCATCGAACTTCTCAGTGTCTGCTGCC 1980
Qy 2244 GGTGCTGGCA CAGGAA CAGATGCCCTCCGCGGT CATGCCACCCAGAGACTGTCCG 2303
Db 1981 GGTGCTGGCA CAGGAA CAGATGCCCTCCGCGGT CATGCCACCCAGAGACTGTCCG 2040
Qy 2304 TGTCTATGCCCCCACTCATGCTCTCTTGGCTTACACCACTCTCCAGCTGTGACC 2363
Db 2041 TGTCTATGCCCCCACTCATGCTCTCTTGGCTTACACCACTCTCTCCAGCTGTGACC 2100

Qy 2364 ACCGATGTCCACACACCCCAACCACTTGTCTCCACAGCTACCCAGCTACGACATCGTCC 2423
Db 2101 ACCGATGTCCACACACCCCAACCACTTGTCTCCACAGCTACCCAGCTACGACATCGTCC 2160
Qy 2424 TGGTCTCCACAGATATCTTCCCACTGACAGCGCGCCGCCACAGAGGCACAGTCCCCAG 2483
Db 2161 TGGTCTCCACAGATATCTTCCCACTGACAGCGCGCGCCGCCACAGAGGCACAGTCCCCAG 2220
Qy 2484 CCACCTCTGCAACTGCGAGCCCTCAGTGACCCCTTTTAAAGCACCCCTGATTTCTACCAATG 2543
Db 2221 CCACCTCTGCAACTGCGAGCCCTCAGTGACCCCTTTTAAAGCACCCCTGATTTCTACCAATG 2280
Qy 2544 CAAACACATCTGGGTCTCGATTTATGACAGAGACTTTTGGACATACGAGGACCCCTCAGAC 2603
Db 2281 CAAACACATCTGGGTCTCGATTTATGACAGAGACTTTTGGACATACGAGGACCCCTCAGAC 2340
Qy 2604 CGAGGAAACCTGCGCCCAACCCCAACACGCTGCTTATGTAACACGCGGAAAGCGGCCCT 2663
Db 2341 CGAGGAAACCTGCGCCCAACCCCAACACGCTGCTTATGTAACACGCGGAAAGCGGCCCT 2400
Qy 2664 GCTGCCCTCCACACACATACACATCAGTCTGATCTACAGCCCTGTTTCGGCGTCAGAG 2723
Db 2401 GCTGCCCTCCACACACATACACATCAGTCTGATCTACAGCCCTGTTTCGGCGTCAGAG 2460
Qy 2724 TCCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAGTAGGGGCGCAGTTTCCAAATTCACC 2783
Db 2461 TCCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAGTAGGGGCGCAGTTTCCAAATTCACC 2520
Qy 2784 CTGTGAGGAGTGAAGCGGATCTGAGCTTCTTGTGACTTAAAGGTCGCGCTTGGGAAT 2843
Db 2521 CTGTGAGGAGTGAAGCGGATCTGAGCTTCTTGTGACTTAAAGGTCGCGCTTGGGAAT 2580
Qy 2844 AAGTTTGTCTTCTGCGCTTTAGCCTTAAAAAAA 2886
Db 2581 AAGTTTGTCTTCTGCGCTTTAGCCTTAAAAAAA 2623

RESULT 11

ADAI8663
ID ADAI8663 standard; cDNA; 2623 BP.
XX AC ADAI8663;
XX XX
XX DT 20-NOV-2003 (first entry)
XX DE Human PRO polynucleotide #32.
XX KW Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; lung;
KW colon; breast; prostate; rectum; cervix; liver; tumour; cancer;
KW glucose uptake; PFA; adipocyte cell; pericyte cell; proteoglycan;
KW cartilage; inner ear utricular supporting cell; cytokine; A-peptide;
KW factor VIIA; endothelial cell.
XX OS Homo sapiens.
XX XX
XX XX US2003054517-A1.
XX XX
XX PD 20-MAR-2003.
XX XX
XX PF 08-MAY-2002; 2002US-00141755.
XX XX
XX PR 31-MAR-1997; 97WO-US005230.
XX PR 12-JUN-1998; 98WO-US012456.
XX PR 14-JUL-1998; 98WO-US014552.
XX PR 28-AUG-1998; 98WO-US017888.
XX PR 10-SEP-1998; 98WO-US018824.
XX PR 14-SEP-1998; 98WO-US019093.
XX PR 14-SEP-1998; 98WO-US019094.
XX PR 14-SEP-1998; 98WO-US019177.
XX PR 16-SEP-1998; 98WO-US019330.
XX PR 17-SEP-1998; 98WO-US019437.

QY 384 GSCCGAGGCGACCTGCCCGCGGCCACAGCGCTACACCTGCTGGGAAACCTCTGCA 443
DB 121 GSCCGAGGCGACCTGCCCGCGGCCACAGCGCTACACCTGCTGGGAAACCTCTGCA 180
QY 444 GCTACGGGCCGGGGCGCTGTATTACGGGCTCATGCGGCTGAGTAAGAAGTACGGACCGGT 503
DB 181 GCTACGGGCCGGGGCGCTGTATTACGGGCTCATGCGGCTGAGTAAGAAGTACGGACCGGT 240
QY 504 GTTACCATCTACCTGAGGACCTGGGGCGCTGTGTGCTGCTGCTGGTGGGACGAGGCTGT 563
DB 241 GTTACCATCTACCTGAGGACCTGGGGCGCTGTGTGCTGCTGCTGGTGGGACGAGGCTGT 300
QY 564 GCGGAGGCGCTGGGAGGTCAGGCTGAGGATTCAGCGCCGCGGAACCGTAGCCATGCT 623
DB 301 GCGGAGGCGCTGGGAGGTCAGGCTGAGGATTCAGCGCCGCGGAACCGTAGCCATGCT 360
QY 624 GGAAGGACTTTTGATGGCCATGGGGTTTTCTTCTCAAACGGGGAGCGGTGGAGGAGCT 683
DB 361 GGAAGGACTTTTGATGGCCATGGGGTTTTCTTCTCAAACGGGGAGCGGTGGAGGAGCT 420
QY 684 GAGGAAGTTTACCATGCTGTCTGCGGACCTGGGCATGGGACGAGAGGCGAGGA 743
DB 421 GAGGAAGTTTACCATGCTGTCTGCGGACCTGGGCATGGGACGAGAGGCGAGGA 480
QY 744 GCTGATCCAGGCGGAGGCGCGGTGTCTGTGTGAGACATTCAGGGGACAGAAAGACGCC 803
DB 481 GCTGATCCAGGCGGAGGCGCGGTGTCTGTGTGAGACATTCAGGGGACAGAAAGACGCC 540
QY 804 ATTCGATCCCTCCCTGCTGCTGCGCCAGGCCACCTCAAACGCTAGTCTGTCTCCCTCTTT 863
DB 541 ATTCGATCCCTCCCTGCTGCTGCGCCAGGCCACCTCAAACGCTAGTCTGTCTCCCTCTTT 600
QY 864 TGGCCTCGCTTCTCTATGAGATTAAGAGTTCAGGCGCGGTGCTCGGGCGAGCTGGTG 923
DB 601 TGGCCTCGCTTCTCTATGAGATTAAGAGTTCAGGCGCGGTGCTCGGGCGAGCTGGTG 660
QY 924 TACCTGCTGGAGTCAGCTCCAGGGGGGTACAGCTACGAGATGTTCTCTGCTGTTCT 983
DB 661 TACCTGCTGGAGTCAGCTCCAGGGGGGTACAGCTACGAGATGTTCTCTGCTGTTCT 720
QY 984 GCGGCCCTTGCAGGCCGCCCAAGCAGTCTCTCCACAGCTCAGCACTTGGCTGCCCTT 1043
DB 721 GCGGCCCTTGCAGGCCGCCCAAGCAGTCTCTCCACAGCTCAGCACTTGGCTGCCCTT 780
QY 1044 CACAGTCGCGAGGTGCAGCAGCACAGGGACCTGGATGTTCTCGGGCCCGCAGCTGA 1103
DB 781 CACAGTCGCGAGGTGCAGCAGCACAGGGACCTGGATGTTCTCGGGCCCGCAGCTGA 840
QY 1104 CTTTGTGATGCTTCTGCTGAAGATGGCAGGAGGAACAAACCCAGGCACAGAAAT 1163
DB 841 CTTTGTGATGCTTCTGCTGAAGATGGCAGGAGGAACAAACCCAGGCACAGAAAT 900
QY 1164 CACCAACAGAACATGCTGATGACAGTCAATTTATTGCTGTTGCTGGGACGATGACGT 1223
DB 901 CACCAACAGAACATGCTGATGACAGTCAATTTATTGCTGTTGCTGGGACGATGACGT 960
QY 1224 CAGCACCGCTCGGCTATACCTCTGCTCTGATGAATAACCTCATGTCCAAAGTG 1283
DB 961 CAGCACCGCTCGGCTATACCTCTGCTCTGATGAATAACCTCATGTCCAAAGTG 1020
QY 1284 GGTACGTGAGGAGCTGAATCGGAGCTGGGGGTGGCCAGGCACCAAGCTTAGGGGACCG 1343
DB 1021 GGTACGTGAGGAGCTGAATCGGAGCTGGGGGTGGCCAGGCACCAAGCTTAGGGGACCG 1080
QY 1344 TACCCGCTCTCTTACACGACGCGGTTCTGATGAGGCGCAGCGGCTGCTGCGCTGGT 1403
DB 1081 TACCCGCTCTCTTACACGACGCGGTTCTGATGAGGCGCAGCGGCTGCTGCGCTGGT 1140
QY 1404 GCCCATGGGAATACCCCGCACCTCATGGGACCAACCGCTTCCGAGGTTACACCTGCC 1463
DB 1141 GCCCATGGGAATACCCCGCACCTCATGGGACCAACCGCTTCCGAGGTTACACCTGCC 1200
QY 1464 CCAGGGCAGGAGGTTCTCCCTCTGCTGCTCCATCTGCTGATGACCCCAACATCTTCAA 1523

DB 1201 CCAGGGCAGGAGGTTCTCCCTCTCTTGGCTCCATCTCTGATGACCCCAACATCTTCAA 1260
QY 1524 GCACCCAGAGAGTTCAACCCAGACGCTTCTCTGATGAGATGACGCTTTCAGGAGCA 1583
DB 1261 GCACCCAGAGAGTTCAACCCAGACGCTTCTCTGATGAGATGACGCTTTCAGGAGCA 1320
QY 1584 TGAGCGGTTCTGCGCTTCTCTCTAGGGAAGGCTGCTGCTTGGAGAGGCGCTGGCAAA 1643
DB 1321 TGAGCGGTTCTGCGCTTCTCTCTAGGGAAGGCTGCTGCTTGGAGAGGCGCTGGCAAA 1380
QY 1644 AGCGAGGCTTCTCTCTTCTTCCACCACTCTCAAGGCTTCTCTCTGGAGAGCCGCTG 1703
DB 1381 AGCGAGGCTTCTCTCTTCTTCCACCACTCTCAAGGCTTCTCTCTGGAGAGCCGCTG 1440
QY 1704 CCGCGCGACACCTCTGAGCTCAAGGCCACGCTGAGTGCCCTTTCAACATTTCCCCAGC 1763
DB 1441 CCGCGCGACACCTCTGAGCTCAAGGCCACGCTGAGTGCCCTTTTCAATTTCCCCAGC 1500
QY 1764 CTTCCAGCTGCAAGTCCGCTCCACTGACCTTCACTCCACAGCGCAGACAGATGAAGAA 1823
DB 1501 CTTCCAGCTGCAAGTCCGCTCCACTGACCTTCACTCCACAGCGCAGACAGATGAAGAA 1560
QY 1824 GGCAACTTGGAAGTGGTGGTCCCGAGACGCTGCTCCAGGCTCAACAGTGCGGCTAGGA 1883
DB 1561 GGCAACTTGGAAGTGGTGGTCCCGAGACGCTGCTCCAGGCTCAACAGTGCGGCTAGGA 1620
QY 1884 CAGGGTTAATGCTCTCAGAGTGTAACCTGACGGCAGCCACATTTACCGCTGCACTGT 1943
DB 1621 CAGGGTTAATGCTCTCAGAGTGTAACCTGACGGCAGCCACATTTACCGCTGCACTGT 1680
QY 1944 TTTCCGGAGTCTGTCCACGGGCCACACGCTCACTGACTCATGTCTGCTGAAGATGACAA 2003
DB 1681 TTTCCGGAGTCTGTCCACGGGCCACACGCTCACTGACTCATGTCTGAAGATGACAA 1740
QY 2004 CCGCACACCCATACAACTAAGGGCCACAAAGAACTGCTGGGTTAGCTTTCCACAG 2063
DB 1741 CCGCACACCCATACAACTAAGGGCCACAAAGAACTGCTGGGTTAGCTTTCCACAG 1800
QY 2064 ACATAAATATAGTCCATCTGCAATCACAGACATAGCCAGGTAAACCCCACTCCCT 2123
DB 1801 ACATAAATATAGTCCATCTGCAATCACAGACATAGCCAGGTAAACCCCACTCCCT 1860
QY 2124 GGATCTGCAGCCACACGCTGGGAGTCTGCTGTGCTACCTTTCAAGCCACAGAAACGGCCA 2183
DB 1861 GGATCTGCAGCCACACGCTGGGAGTCTGCTGTGCTACCTTTCAAGCCACAGAAACGGCCA 1920
QY 2184 CACATGTTACAGCTCAGCGCTCTCCATTCATCGAACTTCTCAGTCTCCCTGCTCCCT 2243
DB 1921 CACATGTTACAGCTCAGCGCTCTCCATTCATCGAACTTCTCAGTCTCCCTGCTCCCT 1980
QY 2244 GGTGCTGCGCACAGGGAACAGATGCCCTCCGGGTCTATGCCACCCAGAGACTGTCG 2303
DB 1981 GGTGCTGCGCACAGGGAACAGATGCCCTCCGGGTCTATGCCACCCAGAGACTGTCG 2040
QY 2304 TGTCTATGGCCCCAACTCATGCTCCCTCTCTTGGCTACCACTCTCCAGGCTGTGACC 2363
DB 2041 TGTCTATGGCCCCAACTCATGCTCCCTCTCTTGGCTACCACTCTCCAGGCTGTGACC 2100
QY 2364 ACCGATGTCACACACCCCAACCACTTGTCCACAGCTACCCAGTACGATCGTC 2423
DB 2101 ACCGATGTCACACACCCCAACCACTTGTCCACAGCTACCCAGTACGATCGTC 2160
QY 2424 TGGCTCCCGAGAGTATCTTCCCACTGAGACAGCCGCCGCCCCACAGAGGACAGTCCCCAG 2483
DB 2161 TGGCTCCCGAGAGTATCTTCCCACTGAGACAGCCGCCGCCCCACAGAGGACAGTCCCCAG 2220
QY 2484 CCACTCTGCAACTGCGAGCCCTCAGTCAACCCCTTTTAAAGACCCCTGATTTCTACAAATG 2543
DB 2221 CCACTCTGCAACTGCGAGCCCTCAGTCAACCCCTTTTAAAGACCCCTGATTTCTACAAATG 2280
QY 2544 CAAGACATCTGGGTTCTGCAATTCACAGAGACTTTGGACATAGGAGACCTCAGAC 2603

Db 2281 CAAACACATCTGGTCTCGGATTATGCACAGAGACTTTGGACATACGAGGACCCCTCAGAC 2340
Qy 2604 CGGAGGACACCTGCCCAACCCACACACGTCCTTATGTAAACACGCTGGAAGCGGCCCT 2663
Db 2341 CGGAGGAACACCTGCCCAACCCACACACGTCCTTATGTAAACACGCTGGAAGCGGCCCT 2400
Qy 2664 GCTGCCCCCTCCACACACACATACACACTCACTGATCTACAGCCCTGTTGCGGCTCAGAG 2723
Db 2401 GCTGCCCCCTCCACACACACATACACACTCACTGATCTACAGCCCTGTTGCGGCTCAGAG 2460
Qy 2724 TCCCCACTAGACCAGTGGAGGGTTAGACACCAAGTAGGGGCCAGTTTCCAATTCAACC 2783
Db 2461 TCCCCACTAGACCAGTGGAGGGTTAGACACCAAGTAGGGGCCAGTTTCCAATTCAACC 2520
Qy 2784 CTGTCAGGAGTGAAGCGGATCTCAGCTTCTTGTGACTTAAGGTCGGGCTTGGGAATT 2843
Db 2521 CTGTCAGGAGTGAAGCGGATCTCAGCTTCTTGTGACTTAAGGTCGGGCTTGGGAATT 2580
Qy 2844 AAAGTTTGTCTTCTGGCCTTTAGCCTAAACCAAAAAAAAAAAAAA 2886
Db 2581 AAAGTTTGTCTTCTGGCCTTTAGCCTAAACCAAAAAAAAAAAAAA 2623

RESULT 12
ADA61286
ID ADA61286 standard; cDNA; 2623 BP.
XX AC ADA61286;
XX DT 20-NOV-2003 (first entry)
XX DE Homo sapiens.
XX KW Human; secreted and transmembrane protein; PRO; gene; ss;
KW tumour necrosis factor alpha release; TNF-alpha release;
KW glucose uptake modulator; FFA uptake modulator;
KW cell proliferation stimulator; cell differentiation stimulator;
KW cell differentiation inhibitor; cytokine release stimulator; tumour;
KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; chromosome mapping; gene mapping;
KW gene therapy; chromosome identification; chromosome marker.
XX OS Novel.
OS human.
OS secreted.
OS and.
OS transmembrane.
OS protein.
OS PRO1906.
OS cDNA.
XX XX
XX US2003049816-A1.
XX PN
XX PD
XX PF 13-MAR-2003.
XX PF 15-APR-2002; 2002US-00123262.
XX XX
PR 31-MAR-1997; 98WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 16-SEP-1998; 98WO-US019177.
PR 17-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98WO-US019437.
PR 29-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 01-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 29-NOV-1999; 99WO-US023089.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.

Db 1261 GCACCAGAAGAGTTCAACCCAGACCGGTTTCTGGATGCAGATGACGGTTTCAGGAACA 1320
Qy 1584 TGAGCGTTCCTGCGCTTCTCTTAGGAAGCGTGTCTGCTTGGAGAGGCGCTTGGCAAA 1643
Db 1321 TGAGCGTTCCTGCGCTTCTCTTAGGAAGCGTGTCTGCTTGGAGAGGCGCTTGGCAAA 1380
Qy 1644 AGCGGAGCTCTTCTCTTCTTCAACCACTCTCAAGCCTTCTCCCTGGAGAGCGCGTG 1703
Db 1381 AGCGGAGCTCTTCTCTTCTTCAACCACTCTCAAGCCTTCTCCCTGGAGAGCGCGTG 1440
Qy 1704 CCOCGCGACACCTTGAGCCTCAAGCCCAACCGTCACTGAGTGGCCTTTCAACATTTCCCCACG 1763
Db 1441 CCOCGCGACACCTTGAGCCTCAAGCCCAACCGTCACTGAGTGGCCTTTCAACATTTCCCCACG 1500
Qy 1764 CTTCCAGCTCAAGTCCGTCCCACTGACCTTCACTCCACACGACGACCCAGATGAAGAA 1823
Db 1501 CTTCCAGCTCAAGTCCGTCCCACTGACCTTCACTCCACACGACGACCCAGATGAAGAA 1560
Qy 1824 GGCAACTTGAAGTGGTGGGTGCCAGGACCGTGCTCCAGCCTCAACAGTGGGATGGA 1883
Db 1561 GGCAACTTGAAGTGGTGGGTGCCAGGACCGTGCTCCAGCCTCAACAGTGGGATGGA 1620
Qy 1884 CAGGGTTAATGCTCCAGAGTGTACACTGACGGCAGCCACATTTACAGCCTGCAGTTGT 1943
Db 1621 CAGGGTTAATGCTCCAGAGTGTACACTGACGGCAGCCACATTTACAGCCTGCAGTTGT 1680
Qy 1944 TTTCCGGAGTCTGCCACGCGCCACACGCTCACTTGACTCATGCTGCTGAAGTGCACAA 2003
Db 1681 TTTCCGGAGTCTGCCACGCGCCACACGCTCACTTGACTCATGCTGCTGAAGTGCACAA 1740
Qy 2004 CCGCACACCCATACACAACTACAAGGGCCACAAAGCACTCTGGTGTAGCTTTCCACAG 2063
Db 1741 CCGCACACCCATACACAACTACAAGGGCCACAAAGCACTCTGGTGTAGCTTTCCACAG 1800
Qy 2064 ACATAAATATAGTCCATCTGCAATCACAAAGCAGATAGCCAGTAAACCCACCACTCCCGCT 2123
Db 1801 ACATAAATATAGTCCATCTGCAATCACAAAGCAGATAGCCAGTAAACCCACCACTCCCGCT 1860
Qy 2124 GGATCTGACGCCACAGTGGAGTCTGGCTGTACCTTCAAGGCCACAGAAACGGCCA 2183
Db 1861 GGATCTGACGCCACAGTGGAGTCTGGCTGTACCTTCAAGGCCACAGAAACGGCCA 1920
Qy 2184 CACATGTTACAGCTTCACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 2243
Db 1921 CACATGTTACAGCTTCACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 1980
Qy 2244 GGTGCTTGGCACAGGAAACAGATGCCCCCTCCGGGGTCAATGCCACCCAGAGACTGTGC 2303
Db 1981 GGTGCTTGGCACAGGAAACAGATGCCCCCTCCGGGGTCAATGCCACCCAGAGACTGTGC 2040
Qy 2304 TGTCTATGGCCCGCACTCATGCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGACC 2363
Db 2041 TGTCTATGGCCCGCACTCATGCTCCCTCTCTTGGCTACACACTCTCCAGCCTGTGACC 2100
Qy 2364 ACCGATGTCACACACCCCAACACATGTTGCCACACAGTACCCACGCTAGACATGCTGC 2423
Db 2101 ACCGATGTCACACACCCCAACACATGTTGCCACACAGTACCCACGCTAACAATGCTCC 2160
Qy 2424 TGGTCTCCCAAGATATCTTCCCACTGAGACAGCGCCGCCCAACAGAGGACAGTCCCCAG 2483
Db 2161 TGGTCTCCCAAGATATCTTCCCACTGAGACAGCGCGGCCCAACAGAGGACAGTCCCCAG 2220
Qy 2484 CCACCTCTGCAACGAGCCTCTGATCAGCCCTTTTAAAGCAGCCTGATTTCTACCAATG 2543
Db 2221 CCACCTCTGCAACGAGCCTCTGATCAGCCCTTTTAAAGCAGCCTGATTTCTACCAATG 2280
Qy 2544 CAACACATCTGGTCTGCGATTATGACAGAGACTTTTGACATACGAGACCCCTCAGAC 2603
Db 2281 CAACACATCTGGTCTGCGATTATGACAGAGACTTTTGACATACGAGACCCCTCAGAC 2340
Qy 2604 CGAGGAAACCTGCCCCAACCCCAACAGCTGTATGTAACCAAGTGGAAAGGGCCCT 2663
Db 2341 CGAGGAAACCTGCCCCAACCCCAACAGCTGTATGTAACCAAGTGGAAAGGGCCCT 2400

Qy 2664 GCTGCCCCCTCCACACACACATACACTCACTCACTGATCTACAGCCCTGTTCGGCGTCAGAG 2723
Db 2401 GCTGCCCCCTCCACACACACATACACTCACTCACTGATCTACAGCCCTGTTCGGCGTCAGAG 2460
Qy 2724 TCCCCACTAGACCCAGTCGGAAGGGGTTAGAGACCAAGTAGGGCCAGTTTCCAATTCACC 2783
Db 2461 TCCCCACTAGACCCAGTCGGAAGGGGTTAGAGACCAAGTAGGGCCAGTTTCCAATTCACC 2520
Qy 2784 CTGTCAAGGAGTGAAGCCGATCTGACCGTTCCTTGTGACTTAAGGCTCCGGCTTGGGAATT 2843
Db 2521 CTGTCAAGGAGTGAAGCCGATCTGACCGTTCCTTGTGACTTAAGGCTCCGGCTTGGGAATT 2580
Qy 2844 AAAGTTTGTCTGCGCTTTAGCCTTAAAAAAA 2886
Db 2581 AAAGTTTGTCTGCGCTTTAGCCTTAAAAAAA 2623

RESULT 13
ADB19071
ID ADB19071 standard; cDNA; 2623 BP.
XX AC ADB19071;
XX DT 20-NOV-2003 (first entry)
XX DE Novel human secreted and transmembrane protein PRO1906 cDNA.
XX KW Human; secreted and transmembrane protein; PRO; gene; ss;
KW Tumour necrosis factor alpha release; TNF-alpha release;
KW glucose uptake modulator; FFA uptake modulator;
KW cell proliferation stimulator; cell differentiation stimulator;
KW cell differentiation inhibitor; cytokin.
XX OS Homo sapiens.
XX PN US2003068796-A1.
XX PD 10-APR-2003.
XX PF 15-APR-2002; 2002US-00123261.
XX PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.

XX The invention describes 305 nucleic acids encoding PRO (secreted and
CC transmembrane) polypeptides (I). (I) is useful for stimulating the
CC release of TNF- α from human blood, for modulating the uptake of
CC glucose or FFA by skeletal muscle cells or adipocyte cells, for
CC stimulating the proliferation or differentiation of chondrocyte cells,
CC for stimulating the proliferation of or gene expression in pericyte
CC cells, for stimulating the release of proteoglycans from cartilage, for
CC stimulating the proliferation of inner ear utricular supporting cells,
CC for stimulating the proliferation of T-lymphocyte cells, for stimulating
CC the release of a cytokine from PMC cells, for inhibiting the binding of
CC A-peptide to factor VIIA, for inhibiting the differentiation of adipocyte
CC cells, for stimulating proliferation of endothelial cells, for detecting
CC the presence of tumour in a mammal. The tumour is lung, colon, breast,
CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes
CC are useful for isolating genomic and cDNA nucleotide sequences or
CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful
CC in assays to identify other proteins or molecules involved in binding
CC interaction. A polynucleotide (II) encoding (I) is useful in chromosome
CC and gene mapping, in generation of antisense RNA and DNA, in the
CC preparation of PRO polypeptide, for generating transgenic animals or
CC knockout animals which in turn are useful in the development and
CC screening of therapeutically useful reagents, in gene therapy, for
CC chromosome identification, as chromosome marker, and for generating
CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.
CC detecting its expression in specific cells, tissues or serum, and for
CC affinity purification of PRO from recombinant cell culture or natural
CC sources. (I) and (II) are useful for tissue typing. This sequence encodes
CC a novel human secreted and transmembrane PRO polypeptide.

SQ Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Query Match 89.0%; Score 2621.4; DB 8; Length 2623;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 264 GCGGAGCGCTGGAGAGGAGGAGCGACCTGCGGAGATGAGGCGACCGGACCTG 323
DB 1 GCGGAGCGCTGGAGAGGAGGAGCGACCTGCGGAGATGAGGCGACCGGACCTG 60
QY 324 GCGGCTGCTGCTGGCGCTGGCGTGTCTCTGCTGTGCTGAGCTGGCGCTGTCCGGGACCG 383
DB 61 GCGGCTGCTGCTGGCGCTGGCGTGTCTCTGCTGTGCTGAGCTGGCGCTGTCCGGGACCG 120
QY 384 GCGGCGAGGCGACCTGCGGCGCGCGCGCGCGCTACCGCTGCGGAGACCTCTGCA 443
DB 121 GCGGCGAGGCGACCTGCGGCGCGCGCGCGCGCTACCGCTGCGGAGACCTCTGCA 180
QY 444 GCTACGCGCGCGCGCGCTGTATTACGGGCTCATCGGCTGAGTAAGAAGTACGGACCGGT 503
DB 181 GCTACGCGCGCGCGCGCTGTATTACGGGCTCATCGGCTGAGTAAGAAGTACGGACCGGT 240
QY 504 GTTCACCATCTACTGCGGACCTTGGCGGCTGTGGTGTCTGTGGGCGAGGAGGTGT 563
DB 241 GTTCACCATCTACTGCGGACCTTGGCGGCTGTGGTGTCTGTGGGCGAGGAGGTGT 300
QY 564 GCGGAGGCGCTGGGAGGTGAGGTGAGGAGTTCAGGCGCGGCGGAGACCTGAGGATGCT 623
DB 301 GCGGAGGCGCTGGGAGGTGAGGTGAGGAGTTCAGGCGCGGCGGAGACCTGAGGATGCT 360
QY 624 GGAAGGACCTTTGATGCCATGGGCTTTCTCTCCAAAGCGGAGCGGTGGAGGAGCT 683
DB 361 GGAAGGACCTTTGATGCCATGGGCTTTCTCTCCAAAGCGGAGCGGTGGAGGAGCT 420
QY 684 GAGGAAGTTTACCATGCTGTCTGCGGGAACCTGGGATGGGGAAGCGAGAGCGAGGA 743
DB 421 GAGGAAGTTTACCATGCTGTCTGCGGGAACCTGGGATGGGGAAGCGAGAGCGAGGA 480
QY 744 GCTGATCCAGCGAGGCGCGGCTGTCTGTTGGAGACATTCAGGGGACAGAGGACGCC 803
DB 481 GCTGATCCAGCGAGGCGCGGCTGTCTGTTGGAGACATTCAGGGGACAGAGGACGCC 540
QY 804 ATTCGATCCCTCCCTGCTGGCTGGCGGCGGCGGACCTCCAAAGCTAGTCTGCTCCCTCTT 863

DB 541 ATTCGATCCCTCCCTGCTGGCGGCGGCGGACCTCCAAAGCTAGTCTGCTCCCTCTCTT 600
QY 864 TGGCGCTCGCTTCTCCCTATGAGGATAGGAGTTCAGGCGCGTGGTCCGGGAGCTGGTG 923
DB 601 TGGCGCTCGCTTCTCCCTATGAGGATAGGAGTTCAGGCGCGTGGTCCGGGAGCTGGTG 660
QY 924 TACCGCTGCTGGAGTCACTCCAGGGGGGTGACACCTACGAGATGTTCTCTCTGTTCTT 983
DB 661 TACCGCTGCTGGAGTCACTCCAGGGGGGTGACACCTACGAGATGTTCTCTCTGTTCTT 720
QY 984 GCGGCGCTGCGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1043
DB 721 GCGGCGCTGCGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 1044 CACAGTCCGCGAGGTGACGACGACGAGGGAACCTGATGCTTCGGGCGCGGCGGCGG 1103
DB 781 CACAGTCCGCGAGGTGACGACGACGAGGGAACCTGATGCTTCGGGCGCGGCGGCGG 840
QY 1104 CTTTGTGATGCTTCTCTGCTGAGATGCGACAGGAGGAAACAAACCCAGGACAGATTT 1163
DB 841 CTTTGTGATGCTTCTCTGCTGAGATGCGACAGGAGGAAACAAACCCAGGACAGATTT 900
QY 1164 CACCAACAGAACATGCTGATGACAGTCAATTTATTTGCTGTTGCTGGGACGATGCGT 1223
DB 901 CACCAACAGAACATGCTGATGACAGTCAATTTATTTGCTGTTGCTGGGACGATGCGT 960
QY 1224 CAGCACCGGTCGCTTACCTCTCTGCTGATGAATACCTCATGTGTCGAAAGTG 1283
DB 961 CAGCACCGGTCGCTTACCTCTCTGCTGATGAATACCTCATGTGTCGAAAGTG 1020
QY 1284 GGTAGCTGAGGAGCTGAATCGGGAGCTGGGGCTGCGGACGACCAAGCTAGGGACCG 1343
DB 1021 GGTAGCTGAGGAGCTGAATCGGGAGCTGGGGCTGCGGACGACCAAGCTAGGGACCG 1080
QY 1344 TACCGCGCTCTCTTACCGAGCGGCTTCTGATGAGGCGGAGCGGCTGCTGCGCTGGT 1403
DB 1081 TACCGCGCTCTCTTACCGAGCGGCTTCTGATGAGGCGGAGCGGCTGCTGCGCTGGT 1140
QY 1404 GCGCATGGGAATACCCCGCACCTCATGCGGACCAACCCGCTTCCGAGGCTACACCTGCC 1463
DB 1141 GCGCATGGGAATACCCCGCACCTCATGCGGACCAACCCGCTTCCGAGGCTACACCTGCC 1200
QY 1464 CAGGCGACGAGGCTTCTCCCGCTCTTGGCTCCATCTGCGATGACCCCAACATCTCAA 1523
DB 1201 CAGGCGACGAGGCTTCTCCCGCTCTTGGCTCCATCTGCGATGACCCCAACATCTCAA 1260
QY 1524 GCACCCAGAGAGTTCACCCAGACCGTTCCTGGATGCGAGTGGACGGTTTCAGGAAGCA 1583
DB 1261 GCACCCAGAGAGTTCACCCAGACCGTTCCTTGGATGCGAGTGGACGGTTTCAGGAAGCA 1320
QY 1584 TGAGGCGTTCCTGCGCTTCTCTTAGGGAAGCGTGTCTGCTTGGAGAGGCGCTGGCAAA 1643
DB 1321 TGAGGCGTTCCTGCGCTTCTCTTAGGGAAGCGTGTCTGCTTGGAGAGGCGCTGGCAAA 1380
QY 1644 AGCGGAGCTTCTCTCTTCTTCCACCATCTCTACAGGCTTCTCCCTGGAGGCGCGTG 1703
DB 1381 AGCGGAGCTTCTCTCTTCTTCCACCATCTCTACAGGCTTCTCCCTGGAGGCGCGTG 1440
QY 1704 CCGCGCGGACACCTGAGCGCTCAAGCCACCGCTCAGTGGGCTTTTCAACATTTCCCGCAGC 1763
DB 1441 CCGCGCGGACACCTGAGCGCTCAAGCCACCGCTCAGTGGGCTTTTCAACATTTCCCGCAGC 1500
QY 1764 CTTTCCAGTGCAGTCCGTCCTCCACTGACCTTCACTCCACCGACGACGATGAAGGAA 1823
DB 1501 CTTTCCAGTGCAGTCCGTCCTCCACTGACCTTCACTCCACCGACGACGATGAAGGAA 1560
QY 1824 GGCACCTTGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1883
DB 1561 GGCACCTTGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1620
QY 1884 CAGGGTTAATGCTCTCCAGAGTGTACCTGAGGCGGACCAATTTACAGCGCTGCAATGTT 1943

Db 1621 CAGGGTTAATGTCTCCAGAGTGTACACTGAGGCGACACATTTTACACGCTGCAATTGT 1680
QY 1944 TTTCGGAGTGTGTCCACGCGCCACAGCTCAGCTGACTGACTGATGCTGTAGATGACAA 2003
Db 1681 TTTCGGAGTGTGTCCACGCGCCACAGCTCAGCTGACTGACTGATGCTGTAGATGACAA 1740
QY 2004 CCGCACACCCCATACACAACTACAAGGGCCACAAAGCAACTGCTGGTTAGCTTTCCACAG 2063
Db 1741 CCGCACACCCCATACACAACTACAAGGGCCACAAAGCAACTGCTGGTTAGCTTTCCACAG 1800
QY 2064 ACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGGTAAACCCACCAACTCCCT 2123
Db 1801 ACATAAATATAGTCCATCTGCAATCACAAGCACATAGCCAGGTAAACCCACCAACTCCCT 1860
QY 2124 GGATCTGAGCCACACAGTGGAGTCTGGCTGTACCTTCACAGCCACAGAAACGGCCA 2183
Db 1861 GGATCTGAGCCACACAGTGGAGTCTGGCTGTACCTTCACAGCCACAGAAACGGCCA 1920
QY 2184 CACATGTTACAGCTCACACGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 2243
Db 1921 CACATGTTACAGCTCACACGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 1980
QY 2244 GGTGCTGGCACAGGGAACAGCATGCCCCCTCCGGGTTCATGCCACCCAGAGACTGTGC 2303
Db 1981 GGTGCTGGCACAGGGAACAGCATGCCCCCTCCGGGTTCATGCCACCCAGAGACTGTGC 2040
QY 2304 TGTCTATGGCCCCAACTCATGCTCCCTCTCTGGCTACACACTCTCCAGCCTGTGACC 2363
Db 2041 TGTCTATGGCCCCAACTCATGCTCCCTCTCTGGCTACACACTCTCCAGCCTGTGACC 2100
QY 2364 ACCGATGTCCACACACCCCAACACATGTTGCCACACAGCTACCCACGTAACGACATCGTCC 2423
Db 2101 ACCGATGTCCACACACCCCAACACATGTTGCCACACAGCTACCCACGTAACGATCGTCC 2160
QY 2424 TGGCTCCCGAGAGTATCTTCCACTGAGACACGCGCCCCACAGAGGACAGTCCCCAG 2483
Db 2161 TGGCTCCCGAGAGTATCTTCCACTGAGACACGCGCCCCACAGAGGACAGTCCCCAG 2220
QY 2484 CCACCTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAAATG 2543
Db 2221 CCACCTCTGCAACTGACGCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAAATG 2280
QY 2544 CAAACACATCTGGGTCTGCGATTATGCACAGAGACTTTGGACATACGAGGACCCCTCAGAC 2603
Db 2281 CAAACACATCTGGGTCTGCGATTATGCACAGAGACTTTGGACATACGAGGACCCCTCAGAC 2340
QY 2604 CGGAGGACACCTGCCCAACCCCAACCGTGTATGTAAACCGTGGAAAGCGGCCCT 2663
Db 2341 CGGAGGACACCTGCCCAACCCCAACCGTGTATGTAAACCGTGGAAAGCGGCCCT 2400
QY 2664 GCTGCCCTCCACACACATACACTCACTGATCTACAGCCCTGTTTCGGGTGAGAG 2723
Db 2401 GCTGCCCTCCACACACATACACTCACTGATCTACAGCCCTGTTTCGGGTGAGAG 2460
QY 2724 TCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGCCAGTTTCCCAATTCAAC 2783
Db 2461 TCCCACTAGACCCAGTGGAGGGTTAGAGACCAAGTAGGGCCAGTTTCCCAATTCAAC 2520
QY 2784 CTGTGAGGAGTACCGGATCTGAGTTCCTTGTGACTTTAAGGTCGGCTTGGGAAT 2843
Db 2521 CTGTGAGGAGTACCGGATCTGAGTTCCTTGTGACTTTAAGGTCGGCTTGGGAAT 2580
QY 2844 AAAGTTTGTCTTGGCTTTAGCCTAAAAAATAAAAAA 2886
Db 2581 AAAGTTTGTCTTGGCTTTAGCCTAAAAAATAAAAAA 2623

Search completed: September 17, 2004, 08:41:28
Job time : 1094.48 secs

This Page Blank (uspto)

Best Local Similarity 99.1%; Pred. No. 1.9e-109; Matches 562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1668 CACCATCTACAGGCTTCTCCTCGAGAGCCGCGCGGACACCTGAGCTCAA 1727

Db 1 CACCATCTACAGGCTTCTCCTCGAGAGCCGCGCGGACACCTGAGCTCAA 60

QY 1728 GCCACCGTCACTGAGGCTTTTCAACATTCCTCCAGCTTCCAGTGCAGTCCGCTCCAC 1787

Db 61 GCCACCGTCACTGAGGCTTTTCAACATTCCTCCAGCTTCCAGTGCAGTCCGCTCCAC 120

QY 1788 TGACCTTCACTCCACGACGACAGATGAGGAGGCACTTGAAGTGTGGTGCC 1847

Db 121 TGACCTTCACTNCAACGAGGACAGATGAGGAGGCACTTGAAGTGTGGTGCC 180

QY 1848 CAGGACGGTCCCTCCAGCTCAACAGTGGGCAATGACAGGTTAATGTCTCCAGAGTGA 1907

Db 181 CAGGACGGTCCCTCCAGCTCAACAGTGGGCAATGACAGGTTAATGTCTCCAGAGTGA 240

QY 1908 CACTGAGGACGACATTTACAGGCTGAGTGTGTTTCCGGAGTCTGTCCAGGGCC 1967

Db 241 CACTGAGGACGACATTTACAGGCTGAGTGTGTTTCCGGAGTCTGTCCAGGGCC 300

QY 1968 ACACGCTCACTTGACTCATGCTGCTAAGATGACACCGCACACCCATACAACTACAA 2027

Db 301 ACACGCTCACTTGACTCATGCTGCTAAGATGACACCGCACACCCATACAACTACAA 360

QY 2028 GGGCCACAAAGCACTGCTGGGTAGCTTTCCACAGACATAATATAGTCCATCTGCAAT 2087

Db 361 GGGCCACAAAGCACTGCTGGGTAGCTTTCCACAGACATAATATAGTCCATCTGCAAT 420

QY 2088 CACAGACATGACGAGTAAACCCAACTCCCTGGATCTGACGCCACACAGTGGAG 2147

Db 421 CACAAGCACATGACGAGTAAACCCAACTCCCTGGATCTGACGCCACACAGTGGAG 480

QY 2148 TCTGGTGTCACTTCAAGGACACAGAAAGGCGCACATGTTACAGCTCACAGCCCC 2207

Db 481 TCTGGTGTCACTTCAAGGACACAGAAAGGCGCACATGTTACAGCTCACAGCCCC 540

QY 2208 TCTCCATTCATGAACTTCTCAGTGTG 2234

Db 541 TCTCCATTCATGAACTTCTCAGTGTG 567

RESULT 2

US-08-277-031B-8

Sequence 8, Application US/08277031B

Patent No. 6620593

GENERAL INFORMATION:

APPLICANT: Hayashi, Koji

APPLICANT: Sakaki, Toshiyuki

APPLICANT: Yabusaki, Yoshiyasu

APPLICANT: Komai, Koichiro

APPLICANT: Kaneko, Hideo

APPLICANT: Nakatsuka, Iwao

TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF

TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING

TITLE OF INVENTION: HUMAN CYTOCHROME P450

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5inch, 1.44MB

COMPUTER: IBM PC

OPERATING SYSTEM: Dos 5.0

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/277,031B

FILING DATE: 19-JULY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP-201120/1993

APPLICATION NUMBER: JP-180246/1993

APPLICATION NUMBER: JP-208279/1993

FILING DATE: 20-07-1993

FILING DATE: 21-07-1993

FILING DATE: 30-07-1993

ATTORNEY/AGENT INFORMATION:

NAME: Raymond C. Stewart

REGISTRATION NUMBER: 21,066

REFERENCE/DOCKET NUMBER: 20-3530P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-205-8000

TELEFAX: 703-205-8050

TELEX:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1485

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-277-031B-8

Query Match 16.1%; Score 475.2; DB 4; Length 1485;

Best Local Similarity 59.6%; Pred. No. 3.2e-91; Indels 6; Gaps 2;

Matches 839; Conservative 0; Mismatches 563;

QY 378 GACCAGGGCCCCGAGGCCACCTGCCCCCGGGCCGCGCTACCACTCTGGGAAACCT 437

Db 78 GAGGAGAGCAAGGGGAGCTGCTCCGGGACCCACCCCATTTGCCCTTATTGGAACCTA 137

QY 438 CTTGAGAGTACGGCCCGGGCGCTGTATTTCAGGGCTCATCGGCTGAGTAAGAGTACGG 497

Db 138 CTTGAGAGTGAACACAGAGCAGATGTACAACTCCCTCATGAAGATCAGTGAGCGCTATGG 197

QY 498 ACCGGTGTTCACCATCTACCTGGGACCTGGCGGCTGTGGTCTGCTGGGCGAGGA 557

Db 198 CCGCGTGTTCACCATCTACCTGGGCGGCGGCGG---GTCTGGTGTCTGTGGGACATGA 254

QY 558 GGCTGTGGCGGAGGCCCTCGGAGGTTCAGGCTGAGGAGTTTCAGCGCGGGGAAACCGTAGC 617

Db 255 TGCGGTTCAGGAGGCTCTGTTGAGCCAGGCTGAGGAGTTTCAGCGCGGCGGAGCAAGC 314

QY 618 GATGCTGGAAGGACHTTTGATGGCCATGGGTTTCTTCTCCAAAGGGGAGCGGTGGAG 677

Db 315 CACCTTCGACTGGGTCTTCAAAGGCTATGGCGTGGTATTTCAGCAACGGGGAGCGCGCAA 374

QY 678 GCAGCTGAGGAAGTTTACCATGCTTCTCGGGGACCTGGGCTGCGGCAAGCGAGAGG 737

Db 375 GCAGCTCGGGCGGCTTCTCATCGCACCTTCGGGAGCTTCGGGGTGGGCAAGCGAGCAT 434

QY 738 CGAGGAGCTGATCCAGCGGAGGCCCGGTTCTGTGTGGAGACATTCAGGGGAGCAGAGG 797

Db 435 CGAGGAGGCGCATCCAGGAGGAGCGGGCTTCTTCATCGACGCGCTCCGGGGCACTGGCGG 494

QY 798 AGCCCCATTCGATCCCTCCCTGCTGTGGCCAGGCCACCTCCAAAGTAGTCTGTCTCCCT 857

Db 495 CGCCAAATATCGATCCCACTTCTTCTGAGCGCGCACAGTCTCCAAATGTCTACAGTCCAT 554

QY 858 CTTCTTTGGCCTCCGCTTCTCTATGAGGATGAAGAGTTTCCAGGCGGTGTCCGGGAGC 917

Db 555 TGCTTTTGGGACCGCTTTGACTATGAAGGACAAAGAGTTCTCTGCTACTGTTCGCGATGAT 614

QY 918 TGGTGTACCTGCTGGGAGTTCAGCTCCAGGGGGGTTCAGACTACAGATGTTCTCTG 977

Db 615 GCTAGGAATCTTCCAGTTCAGGTCACGCTCAACCTCCAGGGGCGAGCTCTATGAGATGTTCTTC 674

QY 978 GTTCTTGGGGCCCTCCAGGGCCCCCAGAGAGGCTCTCCACCAAGTCTGAGTCTGGC 1037

Db 675 GGTGATGAAGAACCTCTCCAGGACCAAGCAAGGCTTTCAGTTGCTGCAAGGGCTGGA 734

QY	508	ACCATCTACCTGGGACCTGTGGCGCCTGTGTGTGGTCTGTGTTTGGGGAAGAGCTGTGGG	567
Db	199	ACGGTACACCTGGGACC---GAGGGCCGGTGTCATGCTGTGTGGAGTAGAGGCCATACGG	255
QY	568	GAGGCCCTTGGGAGTGCAGCTGAGGAGTTACGGCCCGGGGAACGTTAGCGNATGCTGNA	627
Db	256	GAGGCCCTTGTGGACAAGGCTGAGGCCCTTCTCTGGCCGGGGAAAAATGCCCATGTGTCAAC	315
QY	628	GGGACTTTTGATGGCCATGGGGTTTTTCTTCCAAACGGGGAGCGGTGGAGGCAGCTGAGG	687
Db	316	CCAATTCTTCGGGGATATGGTGTGATCTTTGCCAATGGAAACCGCTGGAAAGTGTTCGG	375
QY	688	AAGTTTACATGCTTGTCTTCGGGACCTGGGCATGGGGAAAGCAGAAGCGGAGAGCTG	747
Db	376	CGATTCTCTGTGACCACTATGAGGGACTTCGGGATGGGAAAGCGAGTGTGGAGAGCGG	435
QY	748	ATCCAGGGCGAGGCCGGTGTCTGTGTGAGACATTCACAGGGGACAGAGGACGCCCATTC	807
Db	436	ATTGAGGAGGAGCTCAGTGTCTGATAGAGGAGCTTCGAAATCCAAAGGGGCCCTCATG	495
QY	808	GATCCCTCCCTGCTGCTGGCCCCAGGCCACCTCCAAACGCTAGTCTGCTCCCTCTCTTTGGC	867
Db	496	GACCCCACCTTCTCTCCAGTCCATTACGCCCAACATCATCTGCTCCATGCTCTTTGGA	555
QY	868	CTCGCTTCTCCTATGAGGATAAGGAGTTCCAGGCCGTGGTCCGGGACAGCTGGTGTGATCC	927
Db	556	AAAAGATTCCCACTCAAGATCAAGAGTTCCTGAAGATGCTGAACCTTCTTACCAAGACT	615
QY	928	CTGTGGGAGTGCAGCTCCAGGGGGGTGACAGCTTCCACCACGTCAGCACCTTGGCTGCCTTCA	1047
Db	616	TTTTTCACTCAGCTCTGTATTTGGCCAGCTGTTTGAGCTCTTCTCTGGCTTCTTGAAA	675
QY	988	CCCCTGCCAGSCCCCCAACAGCAGTCTCTCCACCACGTCAGCACCTTGGCTGCCTTCA	1047
Db	676	TACTTTCTCGGGCACACAGSCAAAGTTTACAAAACCTGCAGGAAATCAAATGCTTACATT	735
QY	1048	GTCGGGAGGTGACGACGACCAGGGGAACTGTGANTGTTTGGGGCCGACAGTGAOCTT	1107
Db	736	GGCCACAGTGTGGAGAAGCACCGTGAACCCCTGGA--CCCCAGCGGCCCCCAAGACCTC	792
QY	1108	GTCGATGCTTCGTCTGAAGTGCACAGGAGCAACAAACCCAGGCACAGAAATTCACC	1167
Db	793	ATCGACACTTACCTTGCTCCATCGAAAAAGAGAAATCCAAACGACACAGTGAATTCAGC	852
QY	1168	AACAAGAACATGCTGATGCACAGTCAITTAATTGTGTGTTGCTGGACGATCACGCTCAGC	1227
Db	853	CACCAGAACCTCAACCTCAACACGCTCTCGCTCTTTCTTGTGTCGACCTGAGACCAACAGC	912
QY	1228	ACCAGGTCGGTATACCTCTGCTCTGTATGAAATACCTCATGTGTCGCAAAAGTGGGTA	1287
Db	913	ACCACTCTCGCTACGGCTTCTCTGCTCAATTAACCTCAATACCTCATGTGTGTCAGAGAGCTC	972
QY	1288	CGTGAGGAGCTGAATCGGGAGCTGGGGCTGGCCAGGACACCAACCTTAGGGACCGTACC	1347
Db	973	TACAGGAGATGNAACAGTGAATTGGGCCCATCGCCCTCCAGAGCTTCATGACCGAGCC	1032
QY	1348	CGCTCCCTTTACACCGACGGGTTTCTGCATGAGCGCAGCGGCTGCTGGGGCTGGTGCC	1407
Db	1033	AAAATGCCATACACAGAGGCAGTCACTATGAGATTACAGATTTTCGACCTTCTCCCC	1092
QY	1408	ATGGGAATACCCCGACCTCATCGGAGACACCGCTTCCGAGGGTACACCTGCCCCAG	1467
Db	1093	ATGGGTGTGCCCACTATGTCAACCAACACACAGCTTCCGAGGGTACATCATCCCCAAG	1152
QY	1468	GGACGAGGTTCTPCCCTCTTGGCTCCATCTGTGATGACCCCAACATCTTCAAGCAC	1527
Db	1153	GACACAGAGTATTTCTATCTCTGAGCAGTCTCTCCATGACCCACACTACTTTGAAAAA	1212
QY	1528	CCAGAGAGTTCAACCCAGACCGTTTCTCTGGATGCAGATGGAACGGTTTCAGGAACATGAG	1587
Db	1213	CCAGACGCTTCAATCTGACCACTTTCTGGATGCCAATGGGACACTGAANAAGACTGAA	1272
QY	1588	CGGTTCTCGCCCTTCTCTTAGGGAACGCTGTCTGCTTGGAGAGGCGCTTGGCAAAAGCG	1647

```

Db      1273  GCTTTATCCCTTCTCTCTAGGGAAGCGGATTGTCTTGTGAAGCATCGCCCGTGCG 1332
QY      1648  GAGCTCTTCTCTTCTTACACACCATCGTACAAAGCCTTCTCCCTGGAGAGCCCGTGCCCG 1707
Db      1333  GAAATGTTCTCTTCTTCTTACACCATCTCCAGAACTTCTCCATGGCGCAGCCCGTGCGCC 1392
QY      1708  CCGGACACCTCTGAGCGCTCAAGCCACACCGTCAGTGGCGCTTTTCAACATTTCCCCAGACCTTC 1767
Db      1393  CCAGAGACATCGATCTGCACCCAGGAGTGTGTGGGCAAAATACCCCAACATAC 1452
QY      1768  CAGCTGCAAGTCCGTCCC 1785
Db      1453  CAGATCGGCTTCTCGCC 1470

RESULT 7
US-08-750-703-2
; Sequence 2, Application US/08750703
; Patent No. 5891633
; GENERAL INFORMATION:
; APPLICANT: Gonzalez, Frank J.; Idle, Jeffrey R.
; TITLE OF INVENTION: DEFECTS IN DRUG
; TITLE OF INVENTION: METABOLISM
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan
; STREET: 345 Park Ave.
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,703
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07605
; FILING DATE: 16-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dorothy R. Auth
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4196PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CYP2A13
; LOCATION:
; OTHER INFORMATION:
US-08-750-703-2

```

	Query Match	15.7%;	Score 461.2;	DB 2;	Length 1737;
	Best Local Similarity	58.7%;	Pred. No. 3.1e-88;		
	Matches 854;	Conservative 0;	Mismatches 593;	Indels 7;	Gaps 3;
Qy	332	TGCTGGCGCTGGCGCTCTCTCTGCTGACGTGGCGCTCTCGGGACACAGGCCCGCGAG			391
Db	32	TGCTGGCGCTGCTGACTGTGATGGTCTTGATGTCAGTCTGGCGGCACGAGACGAGG			91
Qy	392	GCCACTGCCCCCGGGCCACGCCCGCTACCTACCTGTGGGAAACCTCTCTGACGTACGGC			451

92	Db	GGAAAGCTGCCTCGGGACCCACCCCAATTGGCCCTTCATTGGAAATACCTCCAGCTCAACA	151
452	Qy	CCGGGGCGCTGTATTACGGGCTCATCGGCTGAGTAAGAAGTACGGACCGGTGTTCCACCA	511
152	Db	CAGACAGATGTAACTCCCTCATGAAGATCACTGAGCGCTATGCGCCCTGTGTTTCAACA	211
512	Qy	TCTACCTGGACACCTGCGCGCTGTGTGTCTGTGGTTGGCGAGGAGGTGTCGCGGAGG	571
212	Db	TTCACCTTGGGGCCCGGGGG---GTCTGTGTGTGTGTGCGACATGATGCGTCAAGGAGG	268
572	Qy	CCCTGGGAGGTACGGCTAGAGGTTCACGCGCCGGGGAACCGTAGCGATGCTCGAAGGGA	631
269	Db	CTCTGGTGGACACGGCTAGAGGTTCACGCGCGGAGCGAGCAGGCGCACCTTCGACTGGC	328
632	Qy	CTTTTGTATGGCATGGGGTTTCTTCTCCAAACGGGGAGCGGTGGAGCGAGCTGAGGAAT	691
329	Db	TCCTCAAAGGCTATGGCGTGGCGTTACAGCAACGGGGAGCGCGCAAGCATCTCCGGCGT	388
692	Qy	TTACATGCTTGCTCTGCGGACCTGGGCATGGGGAACGAGAAAGCGAGAGCTGATCC	751
389	Db	TCTCCATCGCCACCTTAAGGGTTTGGCGTGGCAAGCGCGCATCGAGAAACGATCC	448
752	Qy	AGCGGAGGCGCGGTGTCTGTGTGAGACAATTCAGAGGGAAGAGACGCCAATTCGATC	811
449	Db	AGGAGGAGCGGGCTTCCTCATCGACGCCCTCGGGGCACGCACGGCGCAATATCGATC	508
812	Qy	CCTCGCTGCTGCGCCACCGCACCTCCACGTAGTCTGCTCCCTCTCTTTGGGCTCC	871
509	Db	CCACCTTCTTCTGAGCGCGACAGTCTCCAAATGTCATCAGTCTCATTTCTTTGGGGACC	568
872	Qy	GCTTCTCCTATGAGATAAGCAGTTCCAGGCGGTGCTCCGGCGAGCTGTGGTACCTGTC	931
569	Db	GCTTTGACTATGAGCAAAAGTTCTGTCACTGTTCGCGATGATGCTCGGAGGTTC	628
932	Qy	TGGGAGTCAGTCCACAGGGGGTCCAGACCTACAGAGATGTTCTCTCTGTTTCTGCGGCCC	991
629	Db	AGTTCAACGGAACTCCACGGCGAGCTCTATGAGATGTTCTCTTCGTGATGAAACACC	688
992	Qy	TGCCAGGCCCCCAACAGCAGCTCTCCACACGTCAGCACCTTGGCTGCCTTCACAGTCC	1051
689	Db	TGCCAGGACAAGCAA--CAGGCGCTTTAAGGAGCTGCAAGGGCTGGAGGATTCATCGCCA	747
1052	Qy	GSCAGGTCAGCAGCACACAGGGGAACCTGGATGCTTCGGGCCCCGACGCTGACCTTGTCG	1111
748	Db	AGAAGTGGAGCAAAACAGCGCAGCTGGA---TCCCAATTCCTCCACGGGACTTCATCG	804
1112	Qy	ATGCCCTTCTGCTGAAGATGCGACAGGAGGAACAAAACCCAGGCAAGAAATTCACCAACA	1171
805	Db	ACTCCTTTCTCATCCGATCGAGGAGGAGGAAGAAACCCCAACACAGAGTTCTACTTGA	864
1172	Qy	AGAACATGCTGATGACGTCAATTTATTTGCTGTTTGTGGGACCATGACGGTCAGACCA	1231
865	Db	AGAACCTGGTATGACCAACCTGAACCTTCTTTTGGGGCACTGGAACGCTGAGACCA	924
1232	Qy	CGCTCGGCTATACCTCCTGCTCTGTATGAATAATACCTTCATGTCAAAAGTGGGTACGTG	1291
925	Db	CCCTGCGCTACGGTTTCTCTGTGCTCATGAAGACCCAGAGTGGAGGCCAAGTCCATG	984
1292	Qy	AGGAGCTGAATCGGAGCTGGGGCTGGCCAGGCAACCAAGCCTTAGGGGAACGTTACCGCC	1351
985	Db	AGGAGATTGACAGAGTGATCGCAAGAAACCGGACGCCCAAGTTTGGAGCCGGGCCAAGA	1044
1352	Qy	TCCCTTACACGACCGGTTCTGTGATGAGGCGCAGCGCTCTCTGCGCTGCTGCGCATG	1411
1045	Db	TGCCCTACACAGGCGAGTGTATCCACGAGATCCAAAGATTGGAGACATGCTCCCCATG	1104
1412	Qy	GAATACCCGACCTCATGCGGACCAACCGCTTCCGAGGGGTACACCTGCCCCAGGCA	1471
1105	Db	GTTTGGCCCAAGGGTCAACAGGACACAAGTTTCGGGATTTCTTCTCCTTAAGGGCA	1164
1472	Qy	CGGAGGTCTTCCCCCTCTTGCTCCATCTCTGATGACCCCAACATCTTCAAGCACCCAG	1531

Db	1165	CTGAAGTGTTCCTATGCTGGGCTCGAGCTGAGAGNCCCAGGTTCTTCTCCACCCCC	1222
Qy	1532	AAGAGTTCAACCCAGACCGTTTCCTGATGCA GATGGA CGGTTACGAAAGCATGAGCGT	1591
Db	1225	AGGACTGCAGTCCCCAGCACTTCCTGGATGAGAAGGGCGAGTTTAAAGAGAGTATGCTT	1284
Qy	1592	TCTGCGCCCTCTCCTTTAGGGAAGCGTCTGCGCTTGAGAGGGCCTGGCAAAAGCGGAGC	1651
Db	1285	TTGTGCCCTTTTCCATCGGAAACGGTACTGTTTTTGGAGAAAGGCTTGGCCAGATGGAGC	1344
Qy	1652	TCTTCTCTTCTTCCACCAACCATCCTACAAAGCTTCTCCCTGGAGAGCCCGTGCCCGCGG	1711
Db	1345	TCCTTCTCTCTTCAACCACCATCATGAGAACTTTCGCTTCAAGTCCCTTCAGTCGCCTA	1404
Qy	1712	ACACCTGAGGCTCAAGCCCAACCGTCAGTGGCGCTTTTCAACATTTCCCCAGCGCTTCCAGC	1771
Db	1405	AGGATATCGAGTGTTGCCCCCAACACAGTGGGCTTTGCCACGATCCCAAGAACTACACCA	1464
Qy	1772	TGCAAGTCCGTCCC	1785
Db	1465	TGAGCTTCTTGCCC	1478

RESULT 8
 US-09-023-655-1061
 ; Sequence 1061, Application US/09023655
 ; Patent No. 6607879
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/023,655
 ; FILING DATE: HERewith
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ;
 ; INFORMATION FOR SEQ ID NO: 1061:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1825 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g181357
 ;
 ; US-09-023-655-1061

Query Match 14.6%; Score 429.4; DB 4; Length 1825;

Best Local Similarity 58.3%; Pred. No. 1.7e-81;
Matches 814; Conservative 0; Mismatches 571; Indels 12; Gaps 3;

Qy	391	GGCCACCTGCCCCCGGCCCCACGCCGCTACCACTGCTGGGAAACCTCCTCGCAGCTACGG	450
Db	137	GGAAAGCTGCTCGGAGACCCAGACCCCTCTCAATCTGGGAAACCTGCTGCTCTTTC	196
Qy	451	CCCGGGCGCTGTATTCAGGCTCATCGGCTGATGAAGTACGGACCGGTGTTCAAC	510
Db	197	TCCCAAGACATGCTGACTTCTCTCACTAAGCTAGCAAGAGTATGCTCCATGTACACA	256
Qy	511	ATCTACCTGGACCTCGCGCCCTGTGTGTCTGTGTGGGAGGAGGTGTGGCGGAG	570
Db	257	GTGCACCTGGACCCAGCGCG--GGTGTGTCTCTACGCGGTACCAAGCTGTGAAGGAG	313
Qy	571	GCCCTGGAGGTCAAGCTGAGGAGTTCAGCGCGCGGGAACCGTAGCGATGCTGGAGGG	630
Db	314	GCCTGTGTGACAGGAGGAGGAGTTAGTGGCCGCGTGACTACCTGCGCTTTTCAAC	373
Qy	631	ACTTTGATGGCATGGGTTTCTCTCAACGGGAGCGGTGGAGCAGCTGAGGAAG	690
Db	374	TTTACCAAGGCAATGTCATCGCTTCTCCAGTGGGATCGATGGAAGGTCTTGAGACAG	433
Qy	691	TTTACCATGCTTGTCTGCGGACCTGGGCATGGGGAAGCAGAGCGGAGGAGCTGATC	750
Db	434	TTCTCTATCCAGATTTACCGGAATTCGGGATGGGGAAGAGCAATTGAGGAGCGGAATC	493
Qy	751	CAGCGGAGGCCGGTGTCTGTGTGGAGACATTCAGGGGACAGAAGGACCCCATTCGAT	810
Db	494	CTAGAGAGGCGAGCTTCTGCTGGCGAGCTGCGGAAACTGAGCGGAGCCCTTGGAC	553
Qy	811	CCCTCCCTGCTGTGGCCAGGCCACCTCCAAAGTGTCTGCTCCCTCTCTTTGGCCTC	870
Db	554	CCACGCTTGTGTGAGTCGCTCAGTGTCCAAATATCTGTTCCGTGCTCTTCGCGAGC	613
Qy	871	CGCTTCTCTATGAGATAGGAGTTCACGCGGTGTCCGGCAGCTGTGTGTACCTG	930
Db	614	CGCTTGACTATGATGAGCGTCTGCTCACCATTTATCCGCTTATCAATGACAACTTC	673
Qy	931	CTGGGAGTCAAGTCCCGAGGGGGTTCAGACCTACGAGATGTT---CTCTGGTTCCTGCGG	987
Db	674	CAATCATGAGCAGCCCTCGGCGAGTGTGACATCTACACCCAGATTCGCGAGC	733
Qy	988	CCCTGCGAGGCCCCCAAGAGAGTCTCCACAGTCAGACCTTGGCTGCTTCA	1047
Db	734	CTCTGACTGGTGTGCTGGCGGCCAACAGCATCTTCCAGAACTTCAAGTSCCTGAGA	793
Qy	1048	GTCCGCGAGTGCAGCAGCACAGGGGACCTGGATGCTTCGGGCCCGCGCAGCTGACCTT	1107
Db	794	GACCTATGCCCCACAGCTGCCAGCACACCCAGGCTCGTCTCCC-----CGGAGCTTC	847
Qy	1108	GTGATGCTCTCTGCTGAAGATGGCACAGGAGGAACAAACCCAGGACAGAAATTCACC	1167
Db	848	ATCCAGTGTCTCTACCAAGATGGCAGAGGAGAGGAGCCCATGTAGCCATCTCCAC	907
Qy	1168	AACAAGAACATGCTGATGACAGTCAATTTATGCTGTTTGGGACGATGACGGTCAGC	1227
Db	908	ATGGATACCTGCTGATGACCAACATACCTGCTTTTGGGGCCACCAAGACGGTGAGC	967
Qy	1228	ACCACGGTGGGTATACCTCTGCTCTGTGATGAATACCTCATGTCCAAAGTGGGTA	1287
Db	968	ACCACGCTGCACACGCTTCTGTGGCACTCATGAAGTACCCAAAGTTCAAGCCCGCGTG	1027
Qy	1288	CGTGGAGGTGAATCGGAGCTGGGGGTGGGCCAGGCCAACGCTAGGGACCGTACC	1347
Db	1028	CAGGAGGAGATCGACCTCGTGTGGACGCGCGCGCTGCGGCGCTGAAGACCCGCGG	1087
Qy	1348	CGCTCCCTTACACGACGCGGTTCTGCAATGAGGCGCAGCGGTGCTGGCTGTGGTCC	1407
Db	1088	GCCATGCTTTACACAGACGCGGTGATCCACGAGGTGCAGCGCTTTTCAGACATATCCCC	1147
Qy	1408	ATGGGAATACCCGCACTCATGCGGACCAACCCGCTTCGAGGGGTACACCTGCCCCAG	1467

CLONE: 1749882
US-09-023-655-459

Query Match 11.1%; Score 328; DB 4; Length 341;
Best Local Similarity 99.1%; Pred. No. 2.3e-60;
Matches 339; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1075 AACCTGGATGCTTGGGGCCCCCAGTGACCTTGTGATGCTTCTCTGTAAGATGCA 1134
Db 1 AACCTGGATGCTTGGGGCCCCCAGTGACCTTGTGATGCTTCTCTGTAAGATGCA 60

QY 1135 CAGGAGGACAAACCCAGGACAGAAATTCACCAAGAAATGCTGATGACAGTCATT 1194
Db 61 CAGGAGGACAAACCCAGGACAGAAATTCACCAAGAAATGCTGATGACAGTCATT 120

QY 1195 TATTTGCTGTTTGTCTGGGACGATGACGGTTCAGCACCGGTCGGTATACCTTCCTGCTC 1254
Db 121 TATTTGCTGTTTGTCTGGGACGATGACGGTTCAGCACCGGTCGGTATACCTTCCTGCTC 180

QY 1255 CTGATGAATACCTCATGCTCCAAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGG 1314
Db 181 CTGATGAATACCTCATGCTCCAAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGG 240

QY 1315 GCTGGCCAGCACCAAGCCTAGGGGACCGTACCGGCTCCCTTACACCGACGCGTTCG 1374
Db 241 GCTGGCCAGCACCAAGCCTAGGGGACCGTACCGGCTCCCTTACACCGACGCGTTCG 300

QY 1375 CATGAGGCGACGCGGCTGTGGCGCTGTGGCCATGGGAATA 1416
Db 301 CNTGAGGCGCA-CGGCTGTGGCGCTGTGGCCATGGGAATA 341

RESULT 10
US-08-277-031B-15
; Sequence 15, Application US/08277031B
; Patent No. 6620593
; GENERAL INFORMATION:
; APPLICANT: Hayashi, Koji
; APPLICANT: Sakaki, Toshiyuki
; APPLICANT: Yabusaki, Yoshiyasu
; APPLICANT: Komai, Koichiro
; APPLICANT: Kaneko, Hideo
; APPLICANT: Nakatsuka, Iwao
; TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF
; TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING
; TITLE OF INVENTION: HUMAN CYTOCHROME P450
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5inch, 1.44MB
; COMPUTER: IBM PC
; OPERATING SYSTEM: Dos 5.0
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/277,031B
; FILING DATE: 19-JULY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP-201120/1993
; APPLICATION NUMBER: JP-180246/1993
; APPLICATION NUMBER: JP-208279/1993
; FILING DATE: 20-07-1993
; FILING DATE: 21-07-1993
; FILING DATE: 30-07-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Raymond C. Stewart
; REGISTRATION NUMBER: 21,066

REFERENCE/DOCKET NUMBER: 20-3530P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1473
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-277-031B-15

Query Match 11.0%; Score 324; DB 4; Length 1473;
Best Local Similarity 53.0%; Pred. No. 2.8e-59;
Matches 740; Conservative 0; Mismatches 650; Indels 6; Gaps 2;

QY 389 GAGGCCACCTGCCCCCGGGCCACGCGCTACCTGCTGAGTAAGATACGACCGCTTCA 508
Db 77 GAGGAAAACTCCCTCTGCGCCCACTCTCTCCAGTGATTTGAAAAATATCTACAGATAG 136

QY 449 GCGCCGGGGCGCTGTATTTCAGGGCTCATGCGGCTGATGTAAGATACGACCGCTTCA 508
Db 137 ATATTAAAGATGTCAGCAAAATCCTTAACCAATCTCTCAAAAATCTATGGCCCTGTGTTCA 196

QY 509 CCATCTACCTGGGACCCCTGGCGGCTGTGCTGCTCTGCTGCTGCGAGGAGGCTGTGCGG 568
Db 197 CTCTGTATTTTGGCCCTCGAGCCCA--TGGTGTGCTGCATGGATATGAAGTGTGAAG 253

QY 569 AGGCCCTGGAGGTTCAGGCTGAGGAGTTCAGCGCGCGGGAAACCTGAGGATGCTGGAAG 628
Db 254 AAGCCCTGATTGATCTTGAGAGGAGGATTTTCTGGAAGAGGCCATTTCCCACTGGCTGAAA 313

QY 629 GGACTTTTGTATGGCCATGGGTTTCTCTCCAAAGGGGAGCGGTGAGGAGCTGAGGA 688
Db 314 GAGCTAACAGAGGATTTGGAATCGTTTTCAGCAATGGAAGAGATGGAAGAGATCCGGC 373

QY 689 AGTTTACCATGCTTGTCTGCGGACCTGCGGATGGGAAAGCGAGAGGAGGAGCTGA 748
Db 374 GTTCTCCCTCATGAGCTGCGGMAATTTTGGGATGGGAAAGAGGACATTTAGGACCGTG 433

QY 749 TCAGGCGAGGCGCGGCTGTCTGCTGAGAGACATTCAGGGGACAGAGAGCGCCATTTCG 808
Db 434 TTCAAGAGGAAGCGCGCTGCTGTGAGGAGTTGAGAAAAAACCAAGGCTTCAACCTGTG 493

QY 809 ATCCCTCCCTGCTGCTGCGCCAGGCCACCTCCAACTAGTCTGCTCCCTCTCTTTGGCC 868
Db 494 ATCCCACTTTCATCCTGGGCTGTGCTCCCTGCAATGTGATGCTCATATTATTTCCAGA 553

QY 869 TCCGCTTCTCTATGAGGATAAGGAGTTCCAGGCGCTGTGTCGGGCGAGCTGGTGGTACC 928
Db 554 AAGGTTTCGATTATAAGATCAGCAATTTCTTAACTTGATGGAATAATTTGAATGAAAA 613

QY 929 TGCTGGGAGTCAGCTCCAGGGGGGTTCAGACCTACAGATGTTCTCTGCTGTTCTCGGGC 988
Db 614 TCAGGATTTGAGACACCCCTCGGATCCAGATATGCAATAATTTTCCACATATCATTAAT 673

QY 989 CCCTGCCAGGCCCCCAAGAGCTCTCCACCACTGTCAGACCTTGGCTGCTCTCACAG 1048
Db 674 ATTTCCGGGAACCCATACAAATTACTTAATAACCTTGTCTTTATGGAAGTATATTT 733

QY 1049 TCCGCGAGGTGTCAGCAGCACCGGGGAACCTGATGCTTCCGGGCGCCCGACGTCGCTTG 1108
Db 734 TGGAGAAAGTAAAGAAACACCAAGAAATCGATGGA--CATCAACAAACCTTCGGGACTTTA 790

QY 1109 TCAGTGCCTTCTGCTGAGATGGCAGAGAGAAACAAACCCAGGACAGAGATTCACCA 1168
Db 791 TTGATTTGCTTCTGATCAAAATGGAAGAAAGCAACCAACAGTCTGAAATTTACTA 850

QY 1169 ACAAGAACATGCTGATGACAGTCATTTATTTGCTGTTGCTGGGACGATACCGTTCAGCA 1228
Db 851 TTGAAAACTGGTAATCACTGACGCTCACTTACTTGGAGCTGGGACAGAGACAAAGCA 910

[illegible]


```

QY 989 CCCTGCCAGGCCCCACAGCAGCTCTCCACACGTCAGCACCTTGGTGCCTTCACAG 1048
Db 986 ACTTCCCGGAACTCACAACAATTACTTAAACGTTGCTTTTAAAGAGTTATATT 745
QY 1049 TCCGCGAGGTGAGCAGCAGCAGGGAACCTTGGATGCTTGGGCCCCGACGCTGACCTTG 1108
Db 746 TGGAAAAAGTAAAGAACACCAAGAATCAATGGA--CATGAACAACCTCAGGACTTGA 802
QY 1109 TCGATGCTTCTGCTGAGATGGCAGAGGAGAAACAAACCCAGGCACAGAAITTCACA 1168
Db 803 TTGATTGCTTCTGATGAATTTGGAAGGAAAGACAAACCAACATCTGAAITTTACTA 862
QY 1169 ACAAGAACATGCTGATGACAGTCAITTTTCTGCTGGGACGATGACGGTCAGCA 1228
Db 863 TTGAAGCTTGGAAACACCTGACGTTGCTTTTGGAGCTGGGACAGAGCACAAGCA 922
QY 1229 CAACGCTCGGTATACCTCTCTCTCTGATGAATAACCTCATCTTCCAAAGTGGGTAC 1288
Db 923 CAACCTGAGATATGCTCTCTCTCTCTGCTGAAGCACCAGAGGTACAGCTAAAGTCC 982
QY 1289 GTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAGGCACCAAGCCTAGGGACCGTACC 1348
Db 983 AGGAAGAGATTGAACGTGTGATGGCAGAAACCGAGGCCCTGCATGCAAGACAGAGCC 1042
QY 1349 GCCTCCCTTACACCAAGCGGTTCTGATGAGCGCAGCGGCTGCTGGGCTGGTGCCCA 1408
Db 1043 ACATGCCCTACACAGATGCTGTGTGACAGAGTCCAGAGATACCTTGACCTTCTCCCA 1102
QY 1409 TGGGAATACCCGCAACCTCTATCGGAGCACCCGCTTCGAGGGTACACCTTGCCTCCAGG 1468
Db 1103 CCAGCTGCCCCATGCACTGACATTAATTCAGAACTATCTCATTTCCCAAGG 1162
QY 1469 SCAGGAGGTCTTCCCTCTCTGCTGCTCATCTGATGACCCCAACATCTTCAAGCACC 1528
Db 1163 GCACAAACATATTAATTTCCCTGACTTCTGTCTACATGACAAACAAAGATTTCCCAACC 1222
QY 1529 CAGAAAGATTCAACCCAGACCCGTTTCTGATGAGAGCGGTTTCAGGAAGATGAGG 1588
Db 1223 CAGAGATGTTGACCTCATCTTCTGATGAGGTTGCAATTTTAAAGAAAGTAAT 1282
QY 1589 GCTTCTGCTCTCTCTAGGAAAGGCTGTGCTCTGAGAGGCGCTTGGAAAGCGG 1648
Db 1283 ACTTCATGCTTCTCAGCAGGAAACGAAATTTGTGTGGAGAAAGCCCTGGCGGCGATGG 1342
QY 1649 AGCTCTCTCTCTTCAACCACTCTACAGCTTCTCCTGAGAGCCGCTGCGCGC 1708
Db 1343 AGCTGTTTTTATTCCTGACCTCCATTTTACAGAACTTTAACTGAAATCTCTGGTGTACC 1402
QY 1709 CGGACACCTGAGCCTCAAGCCACCGCTGAGTGGCTTTTCAACATTTCCCCAGCCTTCC 1768
Db 1403 CAAAGAACCTTGACACCACTCCAGTTGTCATGTTTGGCTCTGTGGCGCTTCTTACC 1462
QY 1769 AGCTGCAAGTCCGTCC 1784
Db 1463 AGCTGTGCTTCAATCC 1478

```

```

RESULT 14
PCT-US95-05744-4
; Sequence 4, Application PC/TUS9505744
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMEK-SPARKS, Marjorie
; APPLICANT: DE MORALS, Sonia M.F.
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT
; TITLE OF INVENTION: OF S-MEPHNYTOIN METABOLISM
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California

```

```

; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05744
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,821
; FILING DATE: 06-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/201,118
; FILING DATE: 22-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 15280-192-1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US95-05744-4

```

```

Query Match 11.0%; Score 322.4; DB 5; Length 1854;
Best Local Similarity 52.9%; Pred. No. 6.7e-59;
Matches 739; Conservative 0; Mismatches 651; Indels 6; Gaps 2;

QY 389 GAGGCCACCTGCCCCCGGCGCCAGCGCTACCACTGCTGGGAAACCTTCTGCGACTAC 448
Db 89 GAGGAAAACTCTCTCTGCGCCCACTCTCTCCAGTGTGTTGAAATATCTACAGATAG 148
QY 449 GSCCGCGGCGCTGTATTTCAGGGCTCATGCGGCTCATGTAAGAAGTACGACCGGTTTCA 508
Db 149 GTATTAAAGACATCAGCAAAATCTTAAACAATCTCTCAAGAGTCTATGGCCCTGTGTTCA 208
QY 509 CCATCTACTGCGACCCCTGGCGGCTGTGGTGTCTGTTGGCAGGAGGCTGTGCGGG 568
Db 209 CTCTGATTTTGGCCT--GAAACCCATAGTGTGCTCATGATATGAAGCAGTGAAG 265
QY 569 AGGCCCTGGAGGTCAGGCTGAGGAGTTTCAGCGCGCGGGGAAACCTTAGGATGCTGGAAG 628
Db 266 AAGCCCTGATTGATCTTGAGAGGAGTTTCTGGAAGAGGCAATTTTCCCACTGGCTGAAA 325
QY 629 GGACTTTTGTGCGCATGCGGTTTCTTCTCCAAACCGGAGCGGTGGAGGAGCTGAGGA 688
Db 326 GAGCTAACAGAGGATTTGGAATTTTTCAGCAATGGAAGAAATGGAAGAGATCCGCGC 385
QY 689 AGTTTACCATGCTTCTCTGCGGACCTGGGCATGGGAAAGCGAAGAGCGAGGAGCTGA 748
Db 386 GTTCTCTCTCATGACGCTGCGGAATTTTGGATGGGAGAGGAGCATTTAGAGCCGCTG 445
QY 749 TCCAGCGGAGCGCGGCTGTCTGTTGGAGACATTCAGGGGACAGAAAGACGCCCATTCG 808
Db 446 TTCAAGAGGAGCGCGCTGCTGTTGGAGGAGTTTGGAGGAGTTGAGAAAAACCAAGGCTT 505
QY 809 ATCCCTCTCTGCTGCTGCGCCAGGCACTTCAAGCTAGTCTGCTCTCTCTCTTGTGGCC 868
Db 506 ATCCCACTTTTCTCTGCGGCTGTGCTCTCTGCAATGTGATCTGCTCTCTATTTTCCATA 565
QY 869 TCCGCTTCTCTATGAGGATAGGAGTTCCAGGCCGTGGTCCGCGGAGCTGTGTGTACCC 928

```

```
Db 566 AAGCTTTTGATTAAGATCAGCAATTTCTTAACCTTAATGGAAAGTTGAATGAAACA 625
QY 929 TGTGGAGTCACTCCAGGGGGTTCAGACCTACAGATGTTCTCTGTTCTCTGCGG 988
Db 626 TCAAGATTTTGAGAGCCCTGGATCAGATCTGCAATAATTTTCTCTATCATTTGATT 685
QY 989 CCCTGCCAGGCCCCACAGCAGCTCTCCACACGTCAGCACTTGGCTGCTTCACAG 1048
Db 686 ACTTCCCGGAATCACAACAATTAATCTTAAACGTTGCTTTATGAAAGTTATATT 745
QY 1049 TCCGGAGTGCAGCAGCACCAGGGAACTTGATGCTTCCGGCCCGCACGTCAGCTTG 1108
Db 746 TGGAAAGTAAAGAACACCAAGATCAATGGA--CATGAACAACCTCAGACCTTAA 802
QY 1109 TCGATGCTTCTGCTGAAGATGGCAGAGGAGCAAAACCCAGGACAGAAATTCACCA 1168
Db 803 TTGATGCTTCTGATGAAATGGAGAGGAAAGCACAACCAACCATCTGAAATTTACTA 862
QY 1169 ACAAGACATGCTGATCAGACATTTATTTGCTGTTGCTGGGACGATGACGTCAGCA 1228
Db 863 TTGAAGCTTGGAAACACTGCAAGTTGACTTTGTTGGAGCTGGACAGACAGCAAGCA 922
QY 1229 CACAGCTCGGCTATACCTCTCTGCTCTGATGAAATACCTCTCATGTCCAAAGTGGTAC 1288
Db 923 CAACCTGAGATATGCTCTCTCTCTGCTGAGCACCAGAGGTCACAGCTTAAGTCC 982
QY 1289 GTGAGAGCTGAATCGGAGCTGGGCTGGCGAGCACAAGCCTTAGGGAGCGTACCC 1348
Db 983 AGGAAGAGATGAACGTGTGATGGCAGAAACCGGAGCCCTGATCAAGACAGAGGCC 1042
QY 1349 GCCTCCCTTACAGGAGCGGTTCTGATGAGCGGAGCGGCTGCTGGCTGTTGCCCA 1408
Db 1043 ACATGCCCTTACAGAGTGTGTGGACAGGTCAGAGATACCTTGACCTTCTCCCCA 1102
QY 1409 TGGGAATACCCCGACCTCATCGGACCAACCGCTTCCGAGGTCACACCTGCCCGAGG 1468
Db 1103 CCAGCTGCCCATGAGTACCTGTGACATTAATTCAGAACTATCTCAITCCCAAGG 1162
QY 1469 GCAGGAGTCTTCCCGCTCTGCTGCTCACTGAGGAGGAGGAGGAGGAGGAGGAGG 1528
Db 1163 GCACAACCATTAATTTCCCTGACTTCTGCTGATGAGGAGGAGGAGGAGGAGGAGG 1222
QY 1529 CAGAAGATTCACCCAGAGCGTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1588
Db 1223 CAGAGATGTTTGACCTTCATCTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1282
QY 1589 CGTTCCTGCTCTCTCTAGGGAAGCGTCTCTGCTTGGAGAGGAGGAGGAGGAGGAGG 1648
Db 1283 ACTTCATGCTTCTCAGCAGGAAACGGATTTGTGAGGAGGAGGAGGAGGAGGAGGAG 1342
QY 1649 AGCTCTCTCTCTTCAACCACTCTCAAGCCTTCTCCCTGGAGAGGAGGAGGAGGAGG 1708
Db 1343 AGCTGTTTTTATCTGACCTCCATTTTACAGAACTTTAACTCTGTTGACCTCTGTTGACC 1402
QY 1709 CGCAGACCTTGACCTCAGCCACCGTCTGAGTGGCTTTTCAATTTCCCGAGCTTCC 1768
Db 1403 CAAGAAGCTTGACACCTCCAGTGTCAATGGTTTGGCTCTGTGCGCGCTTCTTACC 1462
QY 1769 AGTCAAGTCCGTC 1784
Db 1463 AGCTGTCTCATTC 1478
```

RESULT 15

```
US-08-194-981E-4
; Sequence 4, Application US/08194981E
; Patent No. 5886157
; GENERAL INFORMATION:
; APPLICANT: GUENGERICH, F. Peter
; APPLICANT: GUO, Zuyu
; APPLICANT: SANDHU, Punam
; APPLICANT: GILLAM, Elizabeth M. J.
```

```
; TITLE OF INVENTION: EXPRESSION AND PURIFICATION OF
; TITLE OF INVENTION: HUMAN
; TITLE OF INVENTION: CYTOCHROME P450
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, 127 Peachtree Street, NE
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,981E
; FILING DATE: February 10, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Selby
; REGISTRATION NUMBER: 38,298
; REFERENCE/DOCKET NUMBER: 22000.0022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1419 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-194-981E-4
```

Query Match 10.9%; Score 321.8; DB 2; Length 1419;

Best Local Similarity 52.9%; Pred. No. 8.2e-59;

Matches 739; Conservative 0; Mismatches 652; Indels 6; Gaps 2;

```
QY 388 CGAGGCCACCTGCCCCCGCCCGCCGCTACCACTCTGGGAAACCTCTCGCAGCTA 447
Db 22 CGAGGAAATCTCCTCTGCCCCCTCTCCAGTGATGGAATATCTTACAGATA 81
QY 448 CGCCCCGGGCGCTGTATTTCAGGCTCATGCGGCTGAGTAAGAGTACGACCGGTGTTTC 507
Db 82 GGTATTAAAGACATCAGCAATCTTAACTCTCAAGGCTCTATGGCCCTGTGTTTC 141
QY 508 ACATCTACCTGGGACCTGGCGGCTGTGGTCTGTGGTCTGTGGTCTGTGGGAGGAGGTGCGG 567
Db 142 ACTCTGTATTTTGGGCT---GAAACCCATAGTGGTCTGTCATGATATGAAGCAGTGAAG 198
QY 568 GAGGCCCTGGAGGTCAGCTGAGGAGTTCAGCGGCGGGGAAACCTAGCATGCTTGGAA 627
Db 199 GAAGCCCTGATTGATCTTGAGAGGAGTTCCTGGAAGAGGCAATTTTCCACCTGGGTGAA 258
QY 628 GGACTCTTTTATGGCCATGGGGTTTTCTTCTCAACGGGAGCGGTGGAGGAGCTGAGG 687
Db 259 AGAGCTAACAGAGGATTTGGAATTTTTCAGCAATGGAAGAAATGGAAGGAGATCCGG 318
QY 688 AGTTTACCATGCTTGTCTGCGGACCTGGGATGGGAGGAGGAGGAGGAGGAGGAGGAGG 747
Db 319 CGTTTCTCCTCATGACGCTCGGAATTTTGGATGGGAGAGGAGGAGGAGGAGGAGGAGG 378
QY 748 ATCCAGCGGAGGCCCGGCTGTCTGGTGGAGACATTCAGGAGGAGGAGGAGGAGGAGGAGG 807
Db 379 GTTCAAGAGAGGAGGCCCGCTGCTTGTGGAGGAGTTCAGAAAAACCAAGGCTTCCACCTGT 438
QY 808 GATCCCTCTCTGCTGTCGCCAGGCGCACTCCCAAGCTAGTCTGCTCCCTCTCTTGGC 867
Db 439 GATCCCACTTTTATCTCTGGGCTGTGCTCCCTGCAATGTGATCTGCTCCATTTATTTCCAT 498
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 14:58:20 ; Search time 1305.36 Seconds
(without alignments)
11383.562 Million cell updates/sec

Title: US-10-669-693-1

Perfect score: 2944

Sequence: 1 tttctctgttcttactc.....aaaaaaaaaaaaaaaaa 2944

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2944	100.0	2944	9	US-09-748-127-1
2	2944	100.0	2944	13	US-10-669-693-1
3	2621.4	89.0	2623	13	US-10-147-493-63
4	2621.4	89.0	2623	13	US-10-145-127-63
5	2621.4	89.0	2623	13	US-10-160-503-63
6	2621.4	89.0	2623	13	US-10-143-118-63
7	2621.4	89.0	2623	13	US-10-144-993-63
8	2621.4	89.0	2623	13	US-10-158-787-63
9	2621.4	89.0	2623	13	US-10-140-024-63
10	2621.4	89.0	2623	13	US-10-140-808-63
11	2621.4	89.0	2623	13	US-10-152-405-63
12	2621.4	89.0	2623	13	US-10-127-852A-63
13	2621.4	89.0	2623	13	US-10-127-900A-63
14	2621.4	89.0	2623	13	US-10-128-685A-63

15	2621.4	89.0	2623	13	US-10-131-820A-63	Sequence 63, Appl
16	2621.4	89.0	2623	13	US-10-142-886-63	Sequence 63, Appl
17	2621.4	89.0	2623	13	US-10-146-728-63	Sequence 63, Appl
18	2621.4	89.0	2623	13	US-10-146-786-63	Sequence 63, Appl
19	2621.4	89.0	2623	13	US-10-147-499-63	Sequence 63, Appl
20	2621.4	89.0	2623	13	US-10-157-798-63	Sequence 63, Appl
21	2621.4	89.0	2623	15	US-10-028-072-63	Sequence 63, Appl
22	2621.4	89.0	2623	15	US-10-121-049-63	Sequence 63, Appl
23	2621.4	89.0	2623	15	US-10-123-904-63	Sequence 63, Appl
24	2621.4	89.0	2623	15	US-10-140-470-63	Sequence 63, Appl
25	2621.4	89.0	2623	15	US-10-175-746-63	Sequence 63, Appl
26	2621.4	89.0	2623	15	US-10-176-918-63	Sequence 63, Appl
27	2621.4	89.0	2623	15	US-10-176-921-63	Sequence 63, Appl
28	2621.4	89.0	2623	15	US-10-137-865-63	Sequence 63, Appl
29	2621.4	89.0	2623	15	US-10-140-474-63	Sequence 63, Appl
30	2621.4	89.0	2623	15	US-10-142-431-63	Sequence 63, Appl
31	2621.4	89.0	2623	15	US-10-143-114-63	Sequence 63, Appl
32	2621.4	89.0	2623	15	US-10-140-002-63	Sequence 63, Appl
33	2621.4	89.0	2623	15	US-10-142-419-63	Sequence 63, Appl
34	2621.4	89.0	2623	15	US-10-123-262-63	Sequence 63, Appl
35	2621.4	89.0	2623	15	US-10-142-423-63	Sequence 63, Appl
36	2621.4	89.0	2623	15	US-10-121-050-63	Sequence 63, Appl
37	2621.4	89.0	2623	15	US-10-141-755-63	Sequence 63, Appl
38	2621.4	89.0	2623	15	US-10-143-032-63	Sequence 63, Appl
39	2621.4	89.0	2623	15	US-10-123-108-63	Sequence 63, Appl
40	2621.4	89.0	2623	15	US-10-123-236-63	Sequence 63, Appl
41	2621.4	89.0	2623	15	US-10-123-261-63	Sequence 63, Appl
42	2621.4	89.0	2623	15	US-10-140-921-63	Sequence 63, Appl
43	2621.4	89.0	2623	15	US-10-140-928-63	Sequence 63, Appl
44	2621.4	89.0	2623	15	US-10-121-045-63	Sequence 63, Appl
45	2621.4	89.0	2623	15	US-10-123-292-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-748-127-1
; Sequence 1, Application US/09748127
; Patent No. US20020076774A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000695
; CURRENT APPLICATION NUMBER: US/09/748,127
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2944
; TYPE: DNA
; ORGANISM: Human
US-09-748-127-1

Query Match	100.0%;	Score 2944;	DB 9;	Length 2944;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2944;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTTCTCTGTTGTTACTTCCTATCCGGGGCCCAAGCGGTGTCCTCGCGCCCAAGC	60	
Db	1	TTTCTCTGTTGTTACTTCCTATCCGGGGCCCAAGCGGTGTCCTCGCGCCCAAGC	60	
QY	61	CCCGGTAACCTGGTGACCTCCGAGACATCCGTTGAGCATGAGTTCCCGACATCAGG	120	
Db	61	CCCGGTAACCTGGTGACCTCCGAGACATCCGTTGAGCATGAGTTCCCGACATCAGG	120	
QY	121	CGGCGCGGTGGTCCGGGAGAAACCCGCGGGGGAGATAGCTGCCAGGAGGCGAGG	180	
Db	121	CGGCGCGGTGGTCCGGGAGAAACCCGCGGGGGAGATAGCTGCCAGGAGGCGAGG	180	

QY	181	GGGCTGGGCTAGCTGCCCGCCCGCGCCCTGAATTCTTGGGAGGAGAGCGCCCGGCTC	240	Db	1261	AAATACCTCATGTCCTCAAAAGTGGTACGTAGGAGCTGAATCGGGAGCTGGGGCTGGC	1320
Db	181	GGGCTGGGCTAGCTGCCCGCCCGCGCCCTGAATTCTTGGGAGGAGAGCGCCCGGCTC	240	QY	1321	CAGGCACCAAGCCTAGGGGACCGGTACCCGGCTCCCTTACACCGACGGGTCTTCATGAG	1380
QY	241	CGGCGCTTAATAGCCAGCGCGCGAGCGCTGGGAGAGAGAAAGCGACCGACCTGCC	300	Db	1321	CAGGCACCAAGCCTAGGGGACCGGTACCCGGCTCCCTTACACCGACGGGTCTTCATGAG	1380
Db	241	CGGCGCTTAATAGCCAGCGCGCGAGCGCTGGGAGAGAGAAAGCGACCGACCTGCC	300	QY	1381	GGGAGCGGCTGTGGCGCTGGTGGCCATCGGGAATACCCCGACCCCTCATCGGACACAC	1440
QY	301	GAGATGAGCGACCGCGACCTCGGCGCTGTGTGGCGCTGGCGCTCTCTCTCTGCTG	360	Db	1381	GGGAGCGGCTGTGGCGCTGGTGGCCATCGGGAATACCCCGACCCCTCATCGGACACAC	1440
Db	301	GAGATGAGCGACCGCGACCTCGGCGCTGTGTGGCGCTGGCGCTCTCTCTCTGCTG	360	QY	1441	CGCTTCCGAGGTACACCTTGCCTCCAGGGCACGGAGGTCTTCCCTCTCTTGGGTCCATC	1500
QY	361	ACGCTGCGCTGTCCGGGACCAAGGCGCGAGGCGACCTGCGCCCGCGGCGCACCGCGTA	420	Db	1441	CGCTTCCGAGGTACACCTTGCCTCCAGGGCACGGAGGTCTTCCCTCTCTTGGGTCCATC	1500
Db	361	ACGCTGCGCTGTCCGGGACCAAGGCGCGAGGCGACCTGCGCCCGCGGCGCACCGCGTA	420	QY	1501	CTGATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGCTTCTGGAT	1560
QY	421	CCACTGTGGGAAACCTCTCAGCTACGCGCGCGGCGCTGTATTCAGGGCTCATGGG	480	Db	1501	CTGATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGCTTCTGGAT	1560
Db	421	CCACTGTGGGAAACCTCTCAGCTACGCGCGCGGCGCTGTATTCAGGGCTCATGGG	480	QY	1561	GCAGATGACGGTTTCAAGAGCATGAGCGCTTCCCTGCTCTCTCTTCTCTTCTTCTCTT	1620
QY	481	CTGAGTAAGAAATACGACCGGCTTCAACCATCTACCTGGGACCTTGGCGGCTGTGGTG	540	Db	1561	GCAGATGACGGTTTCAAGAGCATGAGCGCTTCCCTGCTCTCTCTTCTCTTCTCTTCTCT	1620
Db	481	CTGAGTAAGAAATACGACCGGCTTCAACCATCTACCTGGGACCTTGGCGGCTGTGGTG	540	QY	1621	TGCCTTGGAGAGGCTTGGCAAAAGCGAGCTCTTCTCTCTTCTTCAACCATCTCTCAA	1680
QY	541	GTCTGTGGTGGGAGGAGGCTGTGGGAGGCGCTTGGAGGCTTGGAGGTTTCTCTCC	600	Db	1621	TGCCTTGGAGAGGCTTGGCAAAAGCGAGCTCTTCTCTCTTCTTCAACCATCTCTCAA	1680
Db	541	GTCTGTGGTGGGAGGAGGCTGTGGGAGGCGCTTGGAGGCTTGGAGGTTTCTCTCC	600	QY	1681	GCCTTCTCCCTGGAGAGCGGCTGCGCGGACCCCTGAGCCTCAAGCCACCGCTCAGT	1740
QY	601	GGCGGGGAAACGCTAGCATCTGGAAGGACTTTTGTATGCGCATGGGGTTTCTCTCC	660	Db	1681	GCCTTCTCCCTGGAGAGCGGCTGCGCGGACCCCTGAGCCTCAAGCCACCGCTCAGT	1740
Db	601	GGCGGGGAAACGCTAGCATCTGGAAGGACTTTTGTATGCGCATGGGGTTTCTCTCC	660	QY	1741	GGCCTTTTCAACATTTCCCGCAGCCTTCCAGCTGCAAGTCCGTCCACCTGACCTTCACTCC	1800
QY	661	AACGGGAGCGGTGGAGGAGCTGAGGAGTTTACCATGCTTGTCTCTCGGAGACTGGGC	720	Db	1741	GGCCTTTTCAACATTTCCCGCAGCCTTCCAGCTGCAAGTCCGTCCACCTGACCTTCACTCC	1800
Db	661	AACGGGAGCGGTGGAGGAGCTGAGGAGTTTACCATGCTTGTCTCTCGGAGACTGGGC	720	QY	1801	ACCCGACAGACCATGAAAGGAGCAACTTGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGT	1860
QY	721	ATGGGGAAGCGAAGAGGAGAGCTGATCAGGCGGAGGCGCGTGTCTGTGGAGACA	780	Db	1801	ACCCGACAGACCATGAAAGGAGCAACTTGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGT	1860
Db	721	ATGGGGAAGCGAAGAGGAGAGCTGATCAGGCGGAGGCGCGTGTCTGTGGAGACA	780	QY	1861	CCAGCTCAACAGTGGGATGACAGAGGTTTAAATGTCTCCAGAGTGTACACTCAGGACAG	1920
QY	781	TTCAGGGGACAGAGGAGCGCCATTCGATCCCTGCTGCTGGCGCCAGGCGACCTCC	840	Db	1861	CCAGCTCAACAGTGGGATGACAGAGGTTTAAATGTCTCCAGAGTGTACACTCAGGACAG	1920
Db	781	TTCAGGGGACAGAGGAGCGCCATTCGATCCCTGCTGCTGGCGCCAGGCGACCTCC	840	QY	1921	CACATTTACAGCCTGAGTTGTTTCCGGAGTCTGTCCAGCGCCACACGCTCACTTG	1980
QY	841	AACGTAGTCTGCTCCCTCTCTTTGGGCTCCGCTTCTCTATGAGGATGAGGTTCCAG	900	Db	1921	CACATTTACAGCCTGAGTTGTTTCCGGAGTCTGTCCAGCGCCACACGCTCACTTG	1980
Db	841	AACGTAGTCTGCTCCCTCTCTTTGGGCTCCGCTTCTCTATGAGGATGAGGTTCCAG	900	QY	1981	ACTCATGTCTTAAAGTGCACAAACCGGACACCCATACAACTACAAGGGCCACAAGCA	2040
QY	901	GCGGTGGTCCGGGAGCTGGTGGTACCTGTGGGAGTCACTCCAGGGGGTTCAGACC	960	Db	1981	ACTCATGTCTTAAAGTGCACAAACCGGACACCCATACAACTACAAGGGCCACAAGCA	2040
Db	901	GCGGTGGTCCGGGAGCTGGTGGTACCTGTGGGAGTCACTCCAGGGGGTTCAGACC	960	QY	2041	ACTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	2100
QY	961	TACGAGATGCTCTCTGGTTCCTGGCGCCCTGCCAGGCGCCCAACAAGCAGTCTCTCCAC	1020	Db	2041	ACTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	2100
Db	961	TACGAGATGCTCTCTGGTTCCTGGCGCCCTGCCAGGCGCCCAACAAGCAGTCTCTCCAC	1020	QY	2101	CCAGGTAAACCCACCACTCCCTGGATCTGACGCCACACGCTGGGAGTCTGGGTGTCACC	2160
QY	1021	CACGTGAGACCTTGGCTGCTTCAAGTCCGCGAGGTGAGAGCAGCAGGAGAACCTTG	1080	Db	2101	CCAGGTAAACCCACCACTCCCTGGATCTGACGCCACACGCTGGGAGTCTGGGTGTCACC	2160
Db	1021	CACGTGAGACCTTGGCTGCTTCAAGTCCGCGAGGTGAGAGCAGCAGGAGAACCTTG	1080	QY	2161	TTTCAAGGCCACAGAAACGGGCCACATGTTTCAAGCTCAGACGCTCTCCATTCATCG	2220
QY	1081	GATGCTTGGGCGCCGACGTGACCTTGTGATGCTTCTCTGTAAGATGGCACAGGAG	1140	Db	2161	TTTCAAGGCCACAGAAACGGGCCACATGTTTCAAGCTCAGACGCTCTCCATTCATCG	2220
Db	1081	GATGCTTGGGCGCCGACGTGACCTTGTGATGCTTCTCTGTAAGATGGCACAGGAG	1140	QY	2221	AACTTCTAGTGTCCCTGCTGGGAGGAAACAGGAGGAAACAGGAGGAAACAGGAGGAAAC	2280
QY	1141	GAACAAACCCAGGACAGAAATTCACCAACAGAAATGCTGATGACAGTCAATTTATTTG	1200	Db	2221	AACTTCTAGTGTCCCTGCTGGGAGGAAACAGGAGGAAACAGGAGGAAACAGGAGGAAAC	2280
Db	1141	GAACAAACCCAGGACAGAAATTCACCAACAGAAATGCTGATGACAGTCAATTTATTTG	1200	QY	2281	TGATGCCACCCAGAGACTGTGCTGTATGAGGCGCCAACTCATGCTCTCTCTTGGCTA	2340
QY	1201	CTGTTTGTGGGACGATGACGGTACGACCAACCGGTGGGCTATACCGCTCTCTCTGATG	1260	Db	2281	TGATGCCACCCAGAGACTGTGCTGTATGAGGCGCCAACTCATGCTCTCTCTTGGCTA	2340
Db	1201	CTGTTTGTGGGACGATGACGGTACGACCAACCGGTGGGCTATACCGCTCTCTCTGATG	1260	QY	2341	CACCACTCTCCAGCCTGTGACCCAGGATGTCCAGACCCCGGAGGAGTGGGCTGGC	2400
QY	1261	AAATACCTCATGTCCAAAGTGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGC	1320				

Db	2341	CACCACTCCACGCTGTGACCAACGATGTCCACACACCCCAACCACTTGTCACACA	2400
Qy	2401	GCTACCCACGTACGACATCGCTCTGGCTCCCGACAGTATCTTCCACTGAGACACGCCG	2460
Db	2401	GCTACCCACGTACGACATCGCTCTGGCTCCCGACAGTATCTTCCACTGAGACACGCCG	2460
Qy	2461	CCCACAGAGGCAAGTCCCAGCCACCTCTGCAACTGCGCCCTCAGTCACCCCTTTTT	2520
Db	2461	CCCACAGAGGCAAGTCCCAGCCACCTCTGCAACTGCGCCCTCAGTCACCCCTTTTT	2520
Qy	2521	AAGCACCTGATTTACCAATGCAACACATCTGGGTCTGGGATATGCAAGACATT	2580
Db	2521	AAGCACCTGATTTACCAATGCAACACATCTGGGTCTGGGATATGCAAGACATT	2580
Qy	2581	TGGACATACGAGGACCTTCAGACCGGAGGAACACTGCCCAACCCCAACAGTGCATTATG	2640
Db	2581	TGGACATACGAGGACCTTCAGACCGGAGGAACACTGCCCAACCCCAACAGTGCATTATG	2640
Qy	2641	TAACCACTGGAAAGCGGCCCTGCTGCCCTCCACACACATACACACTCACTCATCT	2700
Db	2641	TAACCACTGGAAAGCGGCCCTGCTGCCCTCCACACACATACACACTCACTCATCT	2700
Qy	2701	ACAGCCCTCTTCCGGGTCAGAGTCCCACTAGACCCAGTGGAGGGTTAGAGACCAAG	2760
Db	2701	ACAGCCCTCTTCCGGGTCAGAGTCCCACTAGACCCAGTGGAGGGTTAGAGACCAAG	2760
Qy	2761	TAGGGGCCATTTCCAATTCAACCTGTGAGGAGTGAGCCGGATCTGAGCTTCCTGTGCA	2820
Db	2761	TAGGGGCCATTTCCAATTCAACCTGTGAGGAGTGAGCCGGATCTGAGCTTCCTGTGCA	2820
Qy	2821	CTTAAAGGTCGGCTTGGGAATTAAGATTGTTCTTGGCTTTAGCCTAAAAA	2880
Db	2821	CTTAAAGGTCGGCTTGGGAATTAAGATTGTTCTTGGCTTTAGCCTAAAAA	2880
Qy	2881	AA	2940
Db	2881	AA	2940
Qy	2941	AAAA 2944	
Db	2941	AAAA 2944	

```

RESULT 2
US-10-669-693-1
; Sequence 1, Application US/10669693
; Publication No. US20040043413A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000685CON
; CURRENT APPLICATION NUMBER: US/10/669,693
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/748,127
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2944
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-669-693-1

Query Match      100.0%; Score 2944; DB 13; Length 2944;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  TTTCCTCTGTTTGCTTACTCCCTATCCGGGGGCCCAAGCGCTGTCTCCGCCGCCCAAGC 60
DB      1  TTTCCTCTGTTTGCTTACTCCCTATCCGGGGGCCCAAGCGCTGTCTCCGCCGCCCAAGC 60

```


Db	961	 CAGCACACCGTGGGTATACCCCTCTGTCTGTGATGAATAACCTCATGTCCAAAGTG	1020
Qy	1284	GGTACGTGAGGAGCTGAATCGGAGCTGGGGGTGGCCAGGCACCAAGCTTAGGGGACCG	1343
Db	1021	GGTACGTGAGGAGCTGATCGGAGCTGGGGGTGGCCAGGCACCAAGCTTAGGGGACCG	1080
Qy	1344	TACCGGCTCCCTTTACACCGACCGGTTTGTGATGAGCGGAGCGGCTGTGGCGCTGGT	1403
Db	1081	TACCGGCTCCCTTTACACCGACCGGTTTGTGATGAGCGGAGCGGCTGTGGCGCTGGT	1140
Qy	1404	GCCCATGGGAATACCCCGACCCCTCATGCGGACCAACCGCTTCCGAGGGTACACCTTCC	1463
Db	1141	GCCCATGGGAATACCCCGACCCCTCATGCGGACCAACCGCTTCCGAGGGTACACCTTCC	1200
Qy	1464	CCAGGACAGGAGTCTTCCCTCTTGGCTCCATCTGCATGACCCCAACATCTTCAA	1523
Db	1201	CCAGGACAGGAGTCTTCCCTCTTGGCTCCATCTGCATGACCCCAACATCTTCAA	1260
Qy	1524	GCACCCAGAGAGTTCAAACCCAGACCGTTTCTTGGATGCAGATGAGCGTTTCAAGA	1583
Db	1261	GCACCCAGAGAGTTCAAACCCAGACCGTTTCTTGGATGCAGATGAGCGTTTCAAGA	1320
Qy	1584	TGAGCGTTCCTGCGCTTCTCTTAGGGAAGCGTGTCTGCTTGGAGAGGGCTGGCAAA	1643
Db	1321	TGAGCGTTCCTGCGCTTCTCTTAGGGAAGCGTGTCTGCTTGGAGAGGGCTGGCAAA	1380
Qy	1644	AGCGGAGTCTTCTCTTCTTCAACCAATCTCTCAAGCCTTCTCCCTGGAGAGCCGTG	1703
Db	1381	AGCGGAGTCTTCTCTTCTTCAACCAATCTCTCAAGCCTTCTCCCTGGAGAGCCGTG	1440
Qy	1704	CCCGCGGACACCTTGAGCTTCAAGCGCAACCGTTCAGTGGCTTTTCAACATTTCCCGC	1763
Db	1441	CCCGCGGACACCTTGAGCTTCAAGCGCAACCGTTCAGTGGCTTTTCAACATTTCCCGC	1500
Qy	1764	CTTCCAGCTGCAAGTTCGTCCACTGCTTCACTCCACCAACGACACAGATGAAGGAA	1823
Db	1501	CTTCCAGCTGCAAGTTCGTCCACTGCTTCACTCCACCAACGACACAGATGAAGGAA	1560
Qy	1824	GGCAATTGGAAGTGGTGGGTGCCAGGACGGTGCTCCAGCCTCAACAGTGGGATGGA	1883
Db	1561	GGCAATTGGAAGTGGTGGGTGCCAGGACGGTGCTCCAGCCTCAACAGTGGGATGGA	1620
Qy	1884	CAGGGTTAATGTCTCCAGATGTACATGAGGAGGACCAATTTCACGCGCTCGAGTTGT	1943
Db	1621	CAGGGTTAATGTCTCCAGATGTACATGAGGAGGACCAATTTCACGCGCTCGAGTTGT	1680
Qy	1944	TTTTCCGGAGTCTGCCACGGCCCCACGGCTCACTTGACTCATGCTGCTAAGATGCACAA	2003
Db	1681	TTTTCCGGAGTCTGTGCCACGGCCCCACGGCTCACTTGACTCATGCTGCTAAGATGCACAA	1740
Qy	2004	CCGCACCCCATACAACTACAAGGGCCACAAAGCAACTGCTGGGTTAGCTTTCACAG	2063
Db	1741	CCGCACCCCATACAACTACAAGGGCCACAAAGCAACTGCTGGGTTAGCTTTCACAG	1800
Qy	2064	ACATAAATATAGTTCATCTGCAATCACAAGCACATAGCCAGGTAACCCACCTCCCT	2123
Db	1801	ACATAAATATAGTTCATCTGCAATCACAAGCACATAGCCAGGTAACCCACCTCCCT	1860
Qy	2124	GGATCTGACGCCACACAGTGGAGTCTGGCTGTACCTTCAAGCCACAGAAACGGCCA	2183
Db	1861	GGATCTGACGCCACACAGTGGAGTCTGGCTGTACCTTCAAGCCACAGAAACGGCCA	1920
Qy	2184	CACATGTTCAAGCTCACAGCCCTCTCATTTATCGAACTTCTCAGTGTCCCTGTCCCT	2243
Db	1921	CACATGTTCAAGCTCACAGCCCTCTTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT	1980
Qy	2244	GGTCCCTGGCACAGGAAACAGCATGCCCTCCGGGGTCATGCCACCCAGAGACTGTGCG	2303
Db	1981	GGTCCCTGGCACAGGAAACAGCATGCCCTCCGGGGTCATGCCACCCAGAGACTGTGCG	2040
Qy	2304	TGTCTATGGCCCCAACTCATGTCTCCCTCTTCTGGGTACACACTCTCCACGCTGTGACC	2363

Db	2041	TGTCATGCGCCCAACTCATGCTCCCTCTCTTGGCTACACCACTCTCCCAGCCTGTGACC	2100
Qy	2364	ACCGATGTCACACACACCCCCAACCACTTGTCACACAGCTTACCACGATGACGACATCGTGCC	2423
Db	2101	ACCGATGTCACACACACCCCCAACCACTTGTCACACAGCTTACCACGATGACCAATCGTGCC	2160
Qy	2424	TGGCTCCCCAGAGTATCTTCCCACTGAGACACCGCGCCCCCAACAGAGGACACATGCTCCCGAG	2483
Db	2161	TGGCTCCCCAGAGTATCTTCCCACTGAGACACCGCGCCCCCAACAGAGGACACATGCTCCCGAG	2220
Qy	2484	CCACCTCTGCAACTGACAGCCCTCAGTCACCCCTTTTAAACACCGCTGATCTTACCAAAATG	2543
Db	2221	CCACCTCTGCAACTGACAGCCCTCAGTCACCCCTTTTAAACACCGCTGATCTTACCAAAATG	2280
Qy	2544	CAAAACATCTGGGCTCTGCGATTATGACAGAGACTTTTGACATACGAGGACCCCTCAGAC	2603
Db	2281	CAAAACATCTGGGCTCTGCGATTATGACAGAGACTTTTGACATACGAGGACCCCTCAGAC	2340
Qy	2604	CGGAGGAACACCTGCCCAAACCCCAACACGTGCTTTATGTAACCAACGTCGGAAGCGGCCCT	2663
Db	2341	CGGAGGACACCTGCCCAAACCCCAACACGTGCTTTATGTAACCAACGTCGGAAGCGGCCCT	2400
Qy	2664	GCTGCCCTCCACACACATACACATCTACTGATCTACAGCCCTCTTTCGGCGCTCAGAG	2723
Db	2401	GCTGCCCTCCACACACATACACATCTACTGATCTACAGCCCTCTTTCGGCGCTCAGAG	2460
Qy	2724	TCCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAAATTCACC	2783
Db	2461	TCCCCACTAGACCCAGTGGAGGGGTTAGAGACCAAGTAGGGGCCAGTTTCCAAATTCACC	2520
Qy	2784	CTGTCTGAGGAGTGAGCCGGATCTCAGCGTTCCTTGTGACTTAAGGGTCCGGCTTGGGAATT	2843
Db	2521	CTGTCTGAGGAGTGAGCCGGATCTCAGCGTTCCTTGTGACTTAAGGGTCCGGCTTGGGAATT	2580
Qy	2844	AAAGTTCTTTCTCGCCCTTTTAGCCTTAAAAAATAAAAAA	2886
Db	2581	AAAGTTCTTTCTCGCCCTTTTAGCCTTAAAAAATAAAAAA	2623

RESULT 7

```

US10-144-993-63
; Sequence 63, Application US/10144993
; Publication NO. US20040038336A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filwaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND THE
; TITLE OF INVENTION: ACIDS ENCODING
; FILE REFERENCE: P33301RC261
; CURRENT APPLICATION NUMBER: US/10/
; CURRENT FILING DATE: 2002-05-13
; Prior Application removed - See File
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-144-993-63

```

Query Match			
Best Local Similarity		89.0%;	Score 2621.4;
Matches 2622;		Conservative	0;
		Mismatches	1;
		Indels	0;
		Gaps	0;
QY	264	GGGGAGCGCTCGGAGAGAGAGAGAGCGGACCTGCGGAGATGGAGGCGAACCGGCACCTG	323
Db	1	GGGGAGCGCTCGGAGAGAGAGAGAGCGGACCTGCGGAGATGGAGGCGAACCGGCACCTG	60
QY	324	GGCGCTGCTGCTGGCGCTGGCGCTGCTCTGCTGCTGAGCTGGCGCTGCTCGGGACGAC	383
Db	61	GGCGCTGCTGCTGGCGCTGGCGCTGCTCTGCTGCTGAGCTGGCGCTGCTCGGGACGAC	120
QY	384	GGCCGAGGCGCACTGCCCCCGGCGCCACGCGCTACCACTGCTGGGAAACCTCTCTGCA	443
Db	121	GGCCGAGGCGCACTGCCCCCGGCGCCACGCGCTACCACTGCTGGGAAACCTCTCTGCA	180
QY	444	GCTACGGCCGGGGCGCTGTATTTCAGGGCTCATGCGGTGAGTAAAGTACGGACCGGT	503
Db	181	GCTACGGCCGGGGCGCTGTATTTCAGGGCTCATGCGGTGAGTAAAGTACGGACCGGT	240
QY	504	GTTCAACCATCTACCTGGGACCTGCGCGCTGTGCTGGTCTGCTGGTGGGAGGAGCTGT	563
Db	241	GTTCAACCATCTACCTGGGACCTGCGCGCTGTGCTGGTCTGCTGGTGGGAGGAGCTGT	300
QY	564	GGGGAGGCGCTGGAGGCTCAGGCTGAGGAGTTACGCGGCGGGGAAACCGTAGCGATGCT	623
Db	301	GGGGAGGCGCTGGAGGCTCAGGCTGAGGAGTTACGCGGCGGGGAAACCGTAGCGATGCT	360
QY	624	GGAAGGACCTTTCATGCGCCATGGGTTTTCTTCTCAAACGGGGAGCGGTGGAGCAGCT	683
Db	361	GGAAGGACCTTTCATGCGCCATGGGTTTTCTTCTCAAACGGGGAGCGGTGGAGCAGCT	420
QY	684	GAGGAAGTTTACCATGCTGCTCTGCGGACCTGGGCATGGGAGCGAGAGCGAGGCGGAG	743
Db	421	GAGGAAGTTTACCATGCTGCTCTGCGGACCTGGGCATGGGAGCGAGAGCGAGGCGGAG	480
QY	744	GCTGATCGAGCGGAGGCGCGGTGTCTGGTGAGACATTCAGGGGACAGAAAGACGCCC	803
Db	481	GCTGATCGAGCGGAGGCGCGGTGTCTGGTGAGACATTCAGGGGACAGAAAGACGCCC	540
QY	804	ATTCGATCCCTCCCTGCTGCGGCCACAGGCACCTCCAAACGTAGTCTGCTCCCTCTT	863
Db	541	ATTCGATCCCTCCCTGCTGCGGCCACAGGCACCTCCAAACGTAGTCTGCTCCCTCTT	600
QY	864	TGGCCTCCGCTTCTCCTATGAGGATAAGGAGTTCCAGGCGCTGTCCGGGAGCTGTGG	923
Db	601	TGGCCTCCGCTTCTCCTATGAGGATAAGGAGTTCCAGGCGCTGTCCGGGAGCTGTGG	660
QY	924	TACCTGCTGGGAGTCACTCCAGGGGGTCCAGACCTACGAGATGTTCTCTGGTTCCT	983
Db	661	TACCTGCTGGGAGTCACTCCAGGGGGTCCAGACCTACGAGATGTTCTCTGGTTCCT	720
QY	984	GGGGCCCTCCAGGGCCCCACAGCAGCTCCCTCCACCACTGACACCTTGGCTGCTT	1043
Db	721	GGGGCCCTCCAGGGCCCCACAGCAGCTCCCTCCACCACTGACACCTTGGCTGCTT	780
QY	1044	CACAGTCCGSCAGGTGCAGCAGCACACAGGGGAACTTGGATGCTTCGGGGCCCCGCA	1103
Db	781	CACAGTCCGSCAGGTGCAGCAGCACACAGGGGAACTTGGATGCTTCGGGGCCCCGCA	840
QY	1104	CCTTGTGATGCTTCTCTGCTGAAGATGGCACAGGAGGAACAAACCCAGGCACAGATT	1163
Db	841	CCTTGTGATGCTTCTCTGCTGAAGATGGCACAGGAGGAACAAACCCAGGCACAGATT	900
QY	1164	CACCAACAAGAACATGCTGATGACAGTCAATTAATTTGCTGTTTGTGGGACGATGCGGT	1223
Db	901	CACCAACAAGAACATGCTGATGACAGTCAATTAATTTGCTGTTTGTGGGACGATGCGGT	960
QY	1224	CAGCACACGGTCCGCTATACCCCTCTGCTGCTGATGAATACCCCTCATGTCCAAAGTG	1283
Db	961	CAGCACACGGTCCGCTATACCCCTCTGCTGATGAATACCCCTCATGTCCAAAGTG	1020

Db 2101 ACCGATGCCACACACCCCAACCACTGTGCCACACAGCTACCCACGCTACAAATCGTCC 2160
QY 2424 TGGTCCCGCAGATCTTCCCACTGAGACACGCGGCCCCACACAGAGCAGTCCCGCAG 2483
Db 2161 TGGTCCCGCAGATCTTCCCACTGAGACACGCGGCCCCACACAGAGCAGTCCCGCAG 2220
QY 2484 CCACCTCTGCAACTGCGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTCAAAATG 2543
Db 2221 CCACCTCTGCAACTGCGCCCTCAGTCAACCCCTTTTAAAGCACCCCTGATTTCAAAATG 2280
QY 2544 CAAACACATCTGGGTCTCGGATTTAGCACAGAGCTTTGGACATACGAGGACCTTCAGAC 2603
Db 2281 CAAACACATCTGGGTCTCGGATTTAGCACAGAGCTTTGGACATACGAGGACCTTCAGAC 2340
QY 2604 CGGAGGACACCTGCCAACCCCAACACGCTGCTTATGTAACACAGCTGGAAGCGGCCCT 2663
Db 2341 CGGAGGACACCTGCCAACCCCAACACGCTGCTTATGTAACACAGCTGGAAGCGGCCCT 2400
QY 2664 GCTGCCCTCCACACACACATACACATCACTGATCTACAGCCCTGTTGGCGTCCAG 2723
Db 2401 GCTGCCCTCCACACACACATACACATCACTGATCTACAGCCCTGTTGGCGTCCAG 2460
QY 2724 TCCCACTAGACCCAGTGGAGGGTTAGACACCAAGTAGGGGCCAGTTTCCAATTAC 2783
Db 2461 TCCCACTAGACCCAGTGGAGGGTTAGACACCAAGTAGGGGCCAGTTTCCAATTAC 2520
QY 2784 CTGTCAGGAGTACCGGATCTGACGTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT 2843
Db 2521 CTGTCAGGAGTACCGGATCTGACGTTCTTGTGACTTAAGGGTCCGGCTTGGGAAT 2580
QY 2844 AAAGTTTGTCTGCGCTTACCTTAAACCAAAAAAAAAAAAAA 2886
Db 2581 AAAGTTTGTCTGCGCTTACCTTAAACCAAAAAAAAAAAAAA 2623

RESULT 8

US-10-158-787-63
; Sequence 63, Application US/10158787
; Publication No. US20040039164A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C449
; CURRENT APPLICATION NUMBER: US/10/158,787
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-158-787-63

Query Match 89.0%; Score 2621.4; DB 13; Length 2623;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 264 CGGAGCGCCTGGGAGGAGAGAGCGGACCTGCCAGATGGAGGCGACCGGACCTG 323
Db 1 CGGAGCGCCTGGGAGGAGAGAGCGGACCTGCCAGATGGAGGCGACCGGACCTG 60
QY 324 GCGGTGCTGCTGGCGCTGGCGCTGCTCTGCTGCTGACGCTGGCGCTGTCGGGACCG 383
Db 61 GCGGTGCTGCTGGCGCTGGCGCTGCTCTGCTGCTGACGCTGGCGCTGTCGGGACCG 120
QY 384 GCGCGAGGCGCACCTGCCCGCCCGGCGCCACGCGCTACACTGCTGGGAAACCTCTGCA 443
Db 121 GCGCGAGGCGCACCTGCCCGCCCGGCGCCACGCGCTACACTGCTGGGAAACCTCTGCA 180
QY 444 GCTACGCGCCCGGGCGCTGTATTACGSGCTCATGCGGCTGAGTAGAGTAGCGACCGT 503
Db 181 GCTACGCGCCCGGGCGCTGTATTACGSGCTCATGCGGCTGAGTAGAGTAGCGACCGT 240
QY 504 GTTCACCATCTACCTGGGACCTGCGCGCTGCTGCTGCTGCTGCTGCTGGGACGAGCT 563
Db 241 GTTCACCATCTACCTGGGACCTGCGCGCTGCTGCTGCTGCTGCTGCTGGGACGAGCT 300
QY 564 GCGGAGGCGCCTGGGAGGTCAGGCTGAGGAGTTACGCGGCGCGGCGGAAACCGTAGCGAT 623
Db 301 GCGGAGGCGCCTGGGAGGTCAGGCTGAGGAGTTACGCGGCGCGGCGGAAACCGTAGCGAT 360
QY 624 GGAAGGACTTTTGTATGGCCATGGGGTTTCTTCTCCACGCGGAGCGCTGGAGGCGAGCT 683
Db 361 GGAAGGACTTTTGTATGGCCATGGGGTTTCTTCTCCACGCGGAGCGCTGGAGGCGAGCT 420
QY 684 GAGGAAGTTTACCATGCTTGTCTCTGCGGACCTGGGACATGGGAGCGAGGAGGAGGA 743
Db 421 GAGGAAGTTTACCATGCTTGTCTCTGCGGACCTGGGACATGGGAGCGAGGAGGAGGA 480
QY 744 GCTGATCCAGCGGAGGCGCGGCTGCTGCTGAGAGCATTTCCAGGCGACAGAGGAGCGCC 803
Db 481 GCTGATCCAGCGGAGGCGCGGCTGCTGCTGAGAGCATTTCCAGGCGCGAGAGGAGCGCC 540
QY 804 ATTGATCCCTCCCTGCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 863
Db 541 ATTGATCCCTCCCTGCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY 864 TGGCGCTCCGCTTCTCTCTATGAGGATTAAGGAGTTTCCAGGCGGCTGGTCCGCGAGCT 923
Db 601 TGGCGCTCCGCTTCTCTCTATGAGGATTAAGGAGTTTCCAGGCGGCTGGTCCGCGAGCT 660
QY 924 TACCTGCTGGGAGTCACTCCAGGGGGTCACTACGAGATGCTTCTCTGTTTCTTCT 983
Db 661 TACCTGCTGGGAGTCACTCCAGGGGGTCACTACGAGATGCTTCTCTGTTTCTTCT 720
QY 984 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1043
Db 721 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780

QY 1044 CACAGTCCGGCAGGTGACAGCAGCAGCGGAACTTGATGCTTCGGCCCGCAGCTGA 1103
Db 781 CACAGTCCGGCAGGTGACAGCAGCAGCGGAACTTGATGCTTCGGCCCGCAGCTGA 840
QY 1104 CTTTGTGATGCTTCCTGCTGAAGATGGCAAGAGGAAACAAACCCAGGCAACAGATT 1163
Db 841 CTTTGTGATGCTTCCTGCTGAAGATGGCAAGAGGAAACAAACCCAGGCAACAGATT 900
QY 1164 CACCAACAAGAACATGCTGATGACAGTCAATTTATTTGCTTTGCTGGGACGATGACGGT 1223
Db 901 CACCAACAAGAACATGCTGATGACAGTCAATTTATTTGCTTTGCTGGGACGATGACGGT 960
QY 1224 CAGCACCAACGCTCGGCTATACCTTCCTGCTCATGAAATACCTCATGTCCTCAAAAGTG 1283
Db 961 CAGCACCAACGCTCGGCTATACCTTCCTGCTCATGAAATACCTCATGTCCTCAAAAGTG 1020
QY 1284 GGTACGTGAGAGTGAAATCGGAGCTGGGGCTGGCCAGGACCAAGCCTTAGGGGACCG 1343
Db 1021 GGTACGTGAGAGTGAAATCGGAGCTGGGGCTGGCCAGGACCAAGCCTTAGGGGACCG 1080
QY 1344 TACCGCCTCCCTTACACCGACGGGTTCTGCATGAGCGGAGCGGCTGCTGGCGTGGT 1403
Db 1081 TACCGCCTCCCTTACACCGACGGGTTCTGCATGAGCGGAGCGGCTGCTGGCGTGGT 1140
QY 1404 GCCATGGGAATATCCCGCACCTTCATGCGGACCAACCGCTTCGAGGGTACACCTGCC 1463
Db 1141 GCCATGGGAATATCCCGCACCTTCATGCGGACCAACCGCTTCGAGGGTACACCTGCC 1200
QY 1464 CAGGGGACGGAGGTCTTCCCTCTCTGCTCCATCTCCATGACCCCAACATCTTCAA 1523
Db 1201 CAGGGGACGGAGGTCTTCCCTCTCTGCTCCATCTCCATGACCCCAACATCTTCAA 1260
QY 1524 GCACCCAGAGAGTCTCAACCCAGACCGTTTCTGATGAGATGGAAGGTTTCAAGAGCA 1583
Db 1261 GCACCCAGAGAGTCTCAACCCAGACCGTTTCTGATGAGATGGAAGGTTTCAAGAGCA 1320
QY 1584 TGAAGGCTTCTCTCTCTTCTAGGAAAGGCTGCTGCTTGGAGAGGCTGCGCAA 1643
Db 1321 TGAAGGCTTCTCTCTCTTCTAGGAAAGGCTGCTGCTTGGAGAGGCTGCGCAA 1380
QY 1644 AGCGAGCTTCTCTCTCTTCTTCAACCACTCTTCAAGCCTTCTCTGAGAGCCGCTG 1703
Db 1381 AGCGAGCTTCTCTCTCTTCTTCAACCACTCTTCAAGCCTTCTCTGAGAGCCGCTG 1440
QY 1704 CCGCCGAGACCTCAGCTCAGCCGCTCAGCCGCTCAGTGGCTTTTCAAATTTCCCGCAG 1763
Db 1441 CCGCCGAGACCTCAGCTCAGCCGCTCAGCCGCTCAGTGGCTTTTCAAATTTCCCGCAG 1500
QY 1764 CTTCCAGCTGCAAGTCCGCTCCCACTGACCTTCACTCCACGACGACGATGAAGGA 1823
Db 1501 CTTCCAGCTGCAAGTCCGCTCCCACTGACCTTCACTCCACGACGACGATGAAGGA 1560
QY 1824 GSCAATTGGAAGTGGTGGTGGCCAGGACGCTGCTCCAGCCTCAACAGTGGGATGGA 1883
Db 1561 GSCAATTGGAAGTGGTGGTGGCCAGGACGCTGCTCCAGCCTCAACAGTGGGATGGA 1620
QY 1884 CAGGGTTAATGCTCCAGAGTGTACACTGAGGAGCCACATTTACCGCTGCGATGTT 1943
Db 1621 CAGGGTTAATGCTCCAGAGTGTACACTGAGGAGCCACATTTACCGCTGCGATGTT 1680
QY 1944 TTTCCGGAGTCTGTCCACGGGCCACACGCTCACTTGCTCATGCTGCTAAGTGCACAA 2003
Db 1681 TTTCCGGAGTCTGTCCACGGGCCACACGCTCACTTGCTCATGCTGCTAAGTGCACAA 1740
QY 2004 CCGCACCCCATACAACTCAAGGGCCACAAAGCACTGCTGGGTAGCTTTCCACAG 2063
Db 1741 CCGCACCCCATACAACTCAAGGGCCACAAAGCACTGCTGGGTAGCTTTCCACAG 1800
QY 2064 ACATAAATATAGTCTGCAATCACAAGACATAGCCAGGTAAACCAACCTCCCT 2123
Db 1801 ACATAAATATAGTCTGCAATCACAAGACATAGCCAGGTAAACCAACCTCCCT 1860

QY 2124 GGATCTGCAGCCCAACACGCTGGAGTCTGGCTGTCACTTCAACAGCCACAGAAACGGCCA 2183
Db 1861 GGATCTGCAGCCCAACACGCTGGAGTCTGGCTGTCACTTCAACAGCCACAGAAACGGCCA 1920
QY 2184 CACATGTTACAGCTCACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 2243
Db 1921 CACATGTTACAGCTCACAGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT 1980
QY 2244 GGTGCTGGCACAGGAAACAGCATGCCCCCTCCGGGTCTAGCCACCCAGAGACTGTGCG 2303
Db 1981 GGTGCTGGCACAGGAAACAGCATGCCCCCTCCGGGTCTAGCCACCCAGAGACTGTGCG 2040
QY 2304 TGTCTATGGCCCCAATCATGCTCCCTCTCTTGGCTACACACTCTCTCCAGCCTGTGACC 2363
Db 2041 TGTCTATGGCCCCAATCATGCTCCCTCTCTTGGCTACACACTCTCTCCAGCCTGTGACC 2100
QY 2364 ACCGATGTCACACACCCCAACACTTGTCCACACAGCTACCCACGTACGACATGCTCC 2423
Db 2101 ACCGATGTCACACACCCCAACACTTGTCCACACAGCTACCCACGTACGACATGCTCC 2160
QY 2424 TGGCTCCCCAGAGTATCTTCCACTGAGACACGCGGCCCCACACAGAGGCAAGTCCCAATG 2483
Db 2161 TGGCTCCCCAGAGTATCTTCCACTGAGACACGCGGCCCCACACAGAGGCAAGTCCCAATG 2220
QY 2484 CCACCTCTGCAACTGCAAGCCCTCAGTCACCCCTTTTAAAGGACCTGATTTCTACCAATG 2543
Db 2221 CCACCTCTGCAACTGCAAGCCCTCAGTCACCCCTTTTAAAGGACCTGATTTCTACCAATG 2280
QY 2544 CAAACACATCTGGGTCTGCGATTTATGACAGAGACTTTTGACATACAGAGCCTCAGAC 2603
Db 2281 CAAACACATCTGGGTCTGCGATTTATGACAGAGACTTTTGACATACAGAGCCTCAGAC 2340
QY 2604 CGAGGAACACCTGCCCCAACCCCAACACGCTGCTTATGTAACCAACGCTGGAAGCGGCCCT 2663
Db 2341 CGAGGAACACCTGCCCCAACCCCAACACGCTGCTTATGTAACCAACGCTGGAAGCGGCCCT 2400
QY 2664 GCTGCCCTCCACACACACATACACATCTACTGATCTACAGCCCTGTTTCGGGTGACAG 2723
Db 2401 GCTGCCCTCCACACACACATACACATCTACTGATCTACAGCCCTGTTTCGGGTGACAG 2460
QY 2724 TCCCCACTAGCCAGTGGAGGGGTAGAGACCAAGTAGGGGCCAGTTTCCAAATTCACC 2783
Db 2461 TCCCCACTAGCCAGTGGAGGGGTAGAGACCAAGTAGGGGCCAGTTTCCAAATTCACC 2520
QY 2784 CTGTAGGAGTGGAGCCGGATCTGAGCTTCTTGTGACTTAAGGCTCCGCTTGGGAAT 2843
Db 2521 CTGTAGGAGTGGAGCCGGATCTGAGCTTCTTGTGACTTAAGGCTCCGCTTGGGAAT 2580
QY 2844 AAAGTTTGTCTTGGCTTTAGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTT 2886
Db 2581 AAAGTTTGTCTTGGCTTTAGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTT 2623

RESULT 9

US-10-140-024-63

; Sequence 63, Application US/10140024

; Publication NO. US20040058424A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gottitsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

:	APPLICANT:	Wood,William			
:	APPLICANT:	Zhang,Zemin			
:	TITLE OF INVENTION:	SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
:	TITLE OF INVENTION:	ACIDS ENCODING THE SAME			
:	FILE REFERENCE:	P3330R1C69			
:	CURRENT APPLICATION NUMBER:	US/10/140,024			
:	CURRENT FILING DATE:	2002-05-06			
:	Prior Application removed - See Palm or File Wrapper				
:	NUMBER OF SEQ ID NOS:	550			
:	SEQ ID NO	63			
:	LENGTH:	2623			
:	TYPE:	DNA			
:	ORGANISM:	Homo Sapien			
:	US-10-140-024-63				
<hr/>					
Query Match		89.0%;	Score	2621.4;	DB 13; Length 2623;
Best Local Similarity		100.0%;	Pred. No.	0;	
Matches 2622;		Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
<hr/>					
QY	264	GCGGAGCGCTGGGAGGAGGAAGGAGCCGACTGCCGAGATGGAGGCACCGCACCTG 323			
DB	1	GCGGAGCGCTGGGAGGAGGAGGAGCCGACTGCCGAGATGGAGGCACCGCACCTG 60			
<hr/>					
QY	324	GGCGTCGTCTGGCGCTGGCGTGTCTCTGTCTGTACGCTGGCGCTCTCGGGACCAG 383			
DB	61	GGCGTCTGTCTGGCGCTGGCGTGTCTCTGTCTGTACGCTGGCGCTCTCGGGACCAG 120			
<hr/>					
QY	384	GGCCCCGAGGCACCTGCCCCCOCGGGCCCGCCGCTACCACTGCTGGAAAACTCCTGCA 443			
DB	121	GGCCCCGAGGCACCTGCCCCCOCGGGCCCGCCGCTACCACTGCTGGAAAACTCCTGCA 180			
<hr/>					
QY	444	GCTACGGCCCCGGGGCGCTGTATTACGGGCTCATCGGGCTTGAGTAAGAAGTACGGACCGGT 503			
DB	181	GCTACGGCCCCGGGGCGCTGTATTACGGGCTCATCGGGCTTGAGTAAGAAGTACGGACCGGT 240			
<hr/>					
QY	504	GTTCCACCATCTACTGGGACCTGTGGCGGCTGTGTGTGTCTGTTGGGACAGAGCTGT 563			
DB	241	GTTCCACCATCTACTGGGACCTGTGGCGGCTGTGTGTGTCTGTTGGGACAGAGCTGT 300			
<hr/>					
QY	564	GCGGGAGGCCCTGGGAGGTACAGGCTGAGGAGTTCAGCGCCGGGGAAACCGTAGCGATGCT 623			
DB	301	GCGGGAGGCCCTGGGAGGTACAGGCTGAGGAGTTCAGCGCCGGGGAAACCGTAGCGATGCT 360			
<hr/>					
QY	624	GGAGGACATTTTATGCCCATCGGGTTTTCTTCTCCAACGGGAGCGGTGGAGGACGCT 683			
DB	361	GGAGGACATTTTATGCCCATCGGGTTTTCTTCTCCAACGGGAGCGGTGGAGGACGCT 420			
<hr/>					
QY	684	GAGGAAGTTTACCATGCTTCTCTCGGGACCTGGGATGGGACGAGGAAGCGCCAGGA 743			
DB	421	GAGGAAGTTTACCATGCTTCTCTCGGGACCTGGGATGGGACGAGGAAGCGCCAGGA 480			
<hr/>					
QY	744	GCTGATCCAGCGGAGGCCCGGCTGTCTGTGTGGAGACATTCACGGGGACAGAAAGACGCC 803			
DB	481	GCTGATCCAGCGGAGGCCCGGCTGTCTGTGTGGAGACATTCACGGGGACAGAAAGACGCC 540			
<hr/>					
QY	804	ATTGATTCCTTCCTGTCTGTGGCCACAGGCACCTCCAAAGTAGTCTGCTCCCTCTCTT 863			
DB	541	ATTGATTCCTTCCTGTCTGTGGCCACAGGCACCTCCAAAGTAGTCTGCTCCCTCTCTT 600			
<hr/>					
QY	864	TGGCTCTCGCTTCTCTATAGGATAAGGAGTTCACGGCCGTGTCTCGGGCAGCTGGTGG 923			
DB	601	TGGCTCTCGCTTCTCTATAGGATAAGGAGTTCACGGCCGTGTCTCGGGCAGCTGGTGG 660			
<hr/>					
QY	924	TACCTGTGGGAGTCAAGTCCCAAGGGGGTTCAGACCTACAGAGTGTTCCTTGGTTCCT 983			
DB	661	TACCTGTGGGAGTCAAGTCCCAAGGGGGTTCAGACCTACAGAGTGTTCCTTGGTTCCT 720			
<hr/>					
QY	984	GCGGCCCTGCGAGGCCCCACAGCAGCTCCTCCACAGCTCAGCACCTTGGCTGCCTT 1043			
DB	721	GCGGCCCTGCGAGGCCCCACAGCAGCTCCTCCACAGCTCAGCACCTTGGCTGCCTT 780			
<hr/>					
QY	1044	CACAGTCCGGCAGGTGCAGCAGCACAGGGGAACTGGATGCTTCGGGCCCGCCAGCTGA 1103			

; CURRENT FILING DATE: 2002-05-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-152-405-63

Query Match 89.0%; Score 2621.4; DB 13; Length 2623;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	264	GGGAGCGCTGGAGAGGAGAGAGCGGACCTGCCAGATGAGGGACCGGACCTG 323
DB	1	GGGAGCGCTGGAGAGGAGAGAGCGGACCTGCCAGATGAGGGACCGGACCTG 60
QY	324	GGCGCTGCTGCTGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 383
DB	61	GGCGCTGCTGCTGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY	384	GGCGCGAGGCGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 443
DB	121	GGCGCGAGGCGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY	444	GCTACGCGCGCGCGCGCTGTATTACGGGCTCATGGGCTGAGTAAAGATACGGACCGGT 503
DB	181	GCTACGCGCGCGCGCGCTGTATTACGGGCTCATGGGCTGAGTAAAGATACGGACCGGT 240
QY	504	GTTACACATCTACTGGGACCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGG 563
DB	241	GTTACACATCTACTGGGACCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGG 300
QY	564	GGGGAGGCGCTGGGAGGTGAGGCTGAGGAGTTTCAGCGGCGGGGAGACCTGATGCT 623
DB	301	GGGGAGGCGCTGGGAGGTGAGGCTGAGGAGTTTCAGCGGCGGGGAGACCTGATGCT 360
QY	624	GGAAGGATTTTATGATGGGATGGGCTTTTCTTCTCCAAAGCGGAGCGGTGGAGGCGAGCT 683
DB	361	GGAAGGATTTTATGATGGGATGGGCTTTTCTTCTCCAAAGCGGAGCGGTGGAGGCGAGCT 420
QY	684	GAGGAATTTTACATGCTTGTCTGCGGACCTGGGACCTGGGACCTGGGACCTGGG 743
DB	421	GAGGAATTTTACATGCTTGTCTGCGGACCTGGGACCTGGGACCTGGGACCTGGG 480
QY	744	GCTGATCCAGGGGAGGCGCGGTGCTGCTGGGAGACATTTCCAGGGACGAGGACGCGCC 803
DB	481	GCTGATCCAGGGGAGGCGCGGTGCTGCTGGGAGACATTTCCAGGGACGAGGACGCGCC 540
QY	804	ATTGATCCCTCCCTGCTGCTGCGCGGCGGACCTCCAAAGCTAGTCTGCTCCCTCTCTT 863
DB	541	ATTGATCCCTCCCTGCTGCTGCGCGGCGGACCTCCAAAGCTAGTCTGCTCCCTCTCTT 600
QY	864	TGGCGCTCCGCTTCTCTATGAGGATAAGAGTTTCCAGCGGCTGGTCCGGCGAGCTGGTG 923
DB	601	TGGCGCTCCGCTTCTCTATGAGGATAAGAGTTTCCAGCGGCTGGTCCGGCGAGCTGGTG 660
QY	924	TACCTGCTGGAGTCAGCTCCAGGGGCTGAGACCTACGAGATGTTCTCTCTGTTCTT 983
DB	661	TACCTGCTGGAGTCAGCTCCAGGGGCTGAGACCTACGAGATGTTCTCTCTGTTCTT 720
QY	984	GGCGGCGCTGCGAGGCGCGGACCAAGAGCTCTCTCCAGCAGTCAGACCTTGGCTGGCTT 1043
DB	721	GGCGGCGCTGCGAGGCGCGGACCAAGAGCTCTCTCCAGCAGTCAGACCTTGGCTGGCTT 780
QY	1044	CACAGTCGGGAGGTGAGAGGACCAAGGAACTTGGATGTTTCGGGCGCGGACCTGGA 1103
DB	781	CACAGTCGGGAGGTGAGAGGACCAAGGAACTTGGATGTTTCGGGCGCGGACCTGGA 840
QY	1104	CCTTGTGATGCTTCTGCTGAAGATGGCAGGAGGAACTTCCAGGACGAGAACTT 1163
DB	841	CCTTGTGATGCTTCTGCTGAAGATGGCAGGAGGAACTTCCAGGACGAGAACTT 900

QY	1164	CACCAACAAGAACATGCTGATGACAGTCAATTAATTTGCTGTTGCTGGACGATGACGGT 1223
DB	901	CACCAACAAGAACATGCTGATGACAGTCAATTAATTTGCTGTTGCTGGACGATGACGGT 960
QY	1224	CAGCACACGCTCGGCTATACCTCTCTGCTCTGATGAAATACCTCTATGTCAAAAGTG 1283
DB	961	CAGCACACGCTCGGCTATACCTCTCTGCTCTGATGAAATACCTCTATGTCAAAAGTG 1020
QY	1284	GGTACGTGAGGAGCTGAATCGGAGCTGGGCGCTGGCGAGCACCAAGCCTTAGGGACCG 1343
DB	1021	GGTACGTGAGGAGCTGAATCGGAGCTGGGCGCTGGCGAGCACCAAGCCTTAGGGACCG 1080
QY	1344	TACCGGCTCTCTTACACGAGCGGTTCTGATGAGCGCAGCGGCTGCTGGCGCTGGT 1403
DB	1081	TACCGGCTCTCTTACACGAGCGGTTCTGATGAGCGCAGCGGCTGCTGGCGCTGGT 1140
QY	1404	GCCCATGGGAATACCCCGCACCTCATGCGAGCACCGCGCTTCGAGGGGTACACCTGCC 1463
DB	1141	GCCCATGGGAATACCCCGCACCTCATGCGAGCACCGCGCTTCGAGGGGTACACCTGCC 1200
QY	1464	CGAGGCGACGAGGTCTTCCCTCTCTTGGCTCCATCTGCTGATGACCCCAACATCTTCAA 1523
DB	1201	CGAGGCGACGAGGTCTTCCCTCTCTTGGCTCCATCTGCTGATGACCCCAACATCTTCAA 1260
QY	1524	GCACCCAGAGAGTTCAACCCAGACCGTTTCTGGATGAGATGAGCGGTTTCAGAAAGCA 1583
DB	1261	GCACCCAGAGAGTTCAACCCAGACCGTTTCTGGATGAGATGAGCGGTTTCAGAAAGCA 1320
QY	1584	TGAGGGCTTCT 1643
DB	1321	TGAGGGCTTCT 1380
QY	1644	AGCGAGCTTCT 1703
DB	1381	AGCGAGCTTCT 1440
QY	1704	CCGCGCGGACACCTGAGGCTCAAGCCCAAGCGGCTCTCTCTCTCTCTCTCTCTCTCTCT 1763
DB	1441	CCGCGCGGACACCTGAGGCTCAAGCCCAAGCGGCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
QY	1764	CTTCCAGCTGCAAGTCCGCTCCACTGACCTTCACTCCACCGAGCGGCTCTCTCTCTCTCT 1823
DB	1501	CTTCCAGCTGCAAGTCCGCTCCACTGACCTTCACTCCACCGAGCGGCTCTCTCTCTCTCT 1560
QY	1824	GGCAACTTGGAAAGTGGTGGGTGCCAGAGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1883
DB	1561	GGCAACTTGGAAAGTGGTGGGTGCCAGAGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1620
QY	1884	CAGGGTTAATGCTCCAGAGTGTACCTGAGGCGGACATTTACACGCTGCACTTGT 1943
DB	1621	CAGGGTTAATGCTCCAGAGTGTACCTGAGGCGGACATTTACACGCTGCACTTGT 1680
QY	1944	TTTCCGAGTCTGTCCCGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2003
DB	1681	TTTCCGAGTCTGTCCCGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
QY	2004	CCGCGACACCCATACAACTACAGGGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCT 2063
DB	1741	CCGCGACACCCATACAACTACAGGGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCT 1800
QY	2064	ACATAAATATAGTCCATCTGCAATCAAGGACATPAGCAGGTAAACCCCAACTCCCCCT 2123
DB	1801	ACATAAATATAGTCCATCTGCAATCAAGGACATPAGCAGGTAAACCCCAACTCCCCCT 1860
QY	2124	GGATCTGAGGCGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2183
DB	1861	GGATCTGAGGCGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1920
QY	2184	CACATGTTTACAGCTTACACGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2243
DB	1921	CACATGTTTACAGCTTACACGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
QY	2244	GGTGGCTGGCACGGGAACAGCATGCCCTCTCCGGGGTCTATGCCACCCAGAGAGTGTGCG 2303

1981 GGTGCTGGCAGAGGAAACAGCATGCCCTCCGGGGTCAATGCCACCGAGAGCTGCGC 2040
2304 TGCTATGAGCCCAACTCATGCTCCCTCTCTTGCTACACCACTCTCCAGCCCTGTGACC 2363
2041 TGTCTATGAGCCCAACTCATGCTCCCTCTCTTGCTACACCACTCTCCAGCCCTGTGACC 2100
2364 ACCGATGTCACACACACCCCAACCACTTGTCCACACAGCTTACCAGTACACATCTGCTCC 2423
2101 ACCGATGTCACACACACCCCAACCACTTGTCCACACAGCTTACCAGTACACATCTGCTCC 2160
2424 TGGTCCCGCAGATATCTTCCACTGACACAGCGCGCCGCCACACAGGACAGTCCCGAG 2483
2161 TGGTCCCGCAGATATCTTCCACTGACACAGCGCGCGCCGCCACACAGGACAGTCCCGAG 2220
2484 CCACCTCTGCAACTGACGACCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2543
2221 CCACCTCTGCAACTGACGACCTCAGTCAACCCCTTTTAAAGCACCTGATTTACCAATG 2280
2544 CAACACATCTGGGTCTCGATATATGACAGAGCTTTGGACATACGAGGACCTTCAGAC 2603
2281 CAACACATCTGGGTCTCGATATATGACAGAGCTTTGGACATACGAGGACCTTCAGAC 2340
2604 CGGAGAACACTGCGCAACCCCAACCACTGCTTATGTAACCACTGGAAGCGGCCCT 2663
2341 CGGAGAACACTGCGCAACCCCAACCACTGCTTATGTAACCACTGGAAGCGGCCCT 2400
2664 GGTGCCCCCACCACACACATACACACTGATCTACAGCCCTGTTGCGCGTCAGAG 2723
2401 GGTGCCCCCACCACACACATACACACTGATCTACAGCCCTGTTGCGCGTCAGAG 2460
2724 TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGCCAGTTTCAATTCACC 2783
2461 TCCCACTAGACCCAGTGAAGGGTTAGAGACCAAGTAGGGCCAGTTTCAATTCACC 2520
2784 CTGTGAGGAGTGAAGGGTCTGAGCTTCTGTTGACTTAAGGGTCCGGCTTGGGAAT 2843
2521 CTGTGAGGAGTGAAGGGTCTGAGCTTCTGTTGACTTAAGGGTCCGGCTTGGGAAT 2580
2844 AAGTTTGTCTTGGCTTTAGCTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2886
2581 AAGTTTGTCTTGGCTTTAGCTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2623

RESULT 12
US-10-127-852A-63
; Sequence 63, Application US/10127852A
; Publication No. US20030203428A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C88
; CURRENT APPLICATION NUMBER: US/10/127,852A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-127-852A-63

Query Match 89.0%; Score 2621.4; DB 13; Length 2623;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2622; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 264 GCGGAGCCCTGGGAGAGGAGGAGCGGACCTCCGAGATGAGGAGGACCGGACCTG 323
Db 1 GCGGAGCCCTGGGAGAGGAGGAGGAGCGGACCTCCGAGATGAGGAGGACCGGACCTG 60

QY 324 GCGGCTGTCTGCGGCTGGCGCTCTCTCTGCTGCTGAGCGTGGCGCTCTCCGGACACAG 383
Db 61 GCGGCTGTCTGCGGCTGGCGCTCTCTCTGCTGCTGAGCGTGGCGCTCTCCGGACACAG 120

QY 384 GCGGCGAGGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 443
Db 121 GCGGCGAGGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180

QY 444 GCTACGGCGCGGCGGCGGCTGTATTTCAGGGCTCATTCAGGGCTGAGTAAAGTACGGACGGT 503
Db 181 GCTACGGCGCGGCGGCGGCTGTATTTCAGGGCTCATTCAGGGCTGAGTAAAGTACGGACGGT 240

QY 504 GTTCACCATCTACCTGGGACCTGGCGGCTGTGGTGTCTCTGGTGGGAGGAGGCTGT 563
Db 241 GTTCACCATCTACCTGGGACCTGGCGGCGGCTGTGGTGTCTCTGGTGGGAGGAGGCTGT 300

QY 564 GCGGAGGCGGCTGGGAGGCTCAGGCTGAGGAGTTCAGCGGCGGCGGCGGCGGCGGCGGCT 623
Db 301 GCGGAGGCGGCTGGGAGGCTCAGGCTGAGGAGTTCAGCGGCGGCGGCGGCGGCGGCGGCT 360

QY 624 GGAAGGAGCTTTTGTATGGGCTATGGGCTTTTCTTCTCAAGCGGAGCGGCTGGGAGGAGCT 683
Db 361 GGAAGGAGCTTTTGTATGGGCTATGGGCTTTTCTTCTCAAGCGGAGCGGCTGGGAGGAGCT 420

QY 684 GAGGAAGTTTACCATCTGCTCTGCGGACCTGGGCTATGGGAGGCGGAGGAGGAGGAGGAG 743
Db 421 GAGGAAGTTTACCATCTGCTCTGCGGAGCTGGGCTATGGGAGGCGGAGGAGGAGGAGGAG 480

QY 744 GCTGATCCAGGCGGAGGCGGCGGCTGTCTGCTGAGAGCATTCAGGAGGAGGAGGAGGAGGCGCC 803
Db 481 GCTGATCCAGGCGGAGGCGGCGGCTGTCTGCTGAGAGCATTCAGGAGGAGGAGGAGGAGGCGCC 540

QY 804 ATTGATCCCTCTCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 863
Db 541 ATTGATCCCTCTCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 600

QY 864 TGCGCTCCGCTCTCTCTATGAGGATAGGAGTTCAGGCGGCGGCTGCTCGGCGGAGCTGTGG 923
Db 601 TGCGCTCCGCTCTCTCTATGAGGATAGGAGTTCAGGCGGCGGCTGCTCGGCGGAGCTGTGG 660

QY 924 TACCTGTGGAGTCAGTCCAGGGGGTCCAGACCTACAGATGTTCTCTGTTCT 983
Db 661 TACCTGTGGAGTCAGTCCAGGGGGTCCAGACCTACAGATGTTCTCTGTTCT 720
QY 984 GGGGCCCCGAGGCCCCCAAGAGCTCTCCACCAAGTCCAGACCTTGGTGCCTT 1043
Db 721 GGGGCCCCGAGGCCCCCAAGAGCTCTCCACCAAGTCCAGACCTTGGTGCCTT 780
QY 1044 CACAGTCGGAGTGCAGCAGCAGGGAACCTGATGTTCTGGGCCCCGAGTGA 1103
Db 781 CACAGTCGGAGTGCAGCAGCAGGGAACCTGATGTTCTGGGCCCCGAGTGA 840
QY 1104 CTTGTGATGCTTCTGCTGAAGATGCAAGAGGAAACAAACCCAGGACAGATTT 1163
Db 841 CTTGTGATGCTTCTGCTGAAGATGCAAGAGGAAACAAACCCAGGACAGATTT 900
QY 1164 CACCAAGAAGACATGCTGATGACAGTCAATTTATTTGCTTGTGAGGAGTACCGT 1223
Db 901 CACCAAGAAGACATGCTGATGACAGTCAATTTATTTGCTTGTGAGGAGTACCGT 960
QY 1224 CAGCACACGCTCGGCTATACCTCCTGCTCTGATGAAATACCTCTATGTCCTGCTG 1283
Db 961 CAGCACACGCTCGGCTATACCTCCTGCTCTGATGAAATACCTCTATGTCCTGCTG 1020
QY 1284 GGTACGTGAGAGTGAATCGGAGCTGGGGGCTGGCCAGGACCAAGCCTAGGGGACG 1343
Db 1021 GGTACGTGAGAGTGAATCGGAGCTGGGGGCTGGCCAGGACCAAGCCTAGGGGACG 1080
QY 1344 TACCGGCTCCCTTACACCGACGCGTCTGATGAGCGCAGCGGCTGCTGGGCTGGT 1403
Db 1081 TACCGGCTCCCTTACACCGACGCGTCTGATGAGCGCAGCGGCTGCTGGGCTGGT 1140
QY 1404 GCCATGGGAATACCCGACCCCTCATGCGGACCAAGCCTAGGGGCTGACCCCTGCC 1463
Db 1141 GCCATGGGAATACCCGACCCCTCATGCGGACCAAGCCTAGGGGCTGACCCCTGCC 1200
QY 1464 CAGGGCAGCGAGTCTTCCCTCTTCCCTTGGCTCCATCTGATGACCCCAACATCTCAA 1523
Db 1201 CAGGGCAGCGAGTCTTCCCTCTTCCCTTGGCTCCATCTGATGACCCCAACATCTCAA 1260
QY 1524 GCACCCAGAGATTCACCCAGACCGTCTTCTGATGAGAGCGTTCAGGAAGCA 1583
Db 1261 GCACCCAGAGATTCACCCAGACCGTCTTCTGATGAGAGCGTTCAGGAAGCA 1320
QY 1584 TGAGGCTTCTGCTTCTTCTAGGAGCGTCTGCTGAGGAGCGCTGAGGAGCGTGGCAA 1643
Db 1321 TGAGGCTTCTGCTTCTTCTAGGAGCGTCTGCTGAGGAGCGCTGAGGAGCGTGGCAA 1380
QY 1644 AGCGAGCTTCTTCTTCTTCAACCACTCTTCAAGCGTCTTCTGAGAGCGCGTG 1703
Db 1381 AGCGAGCTTCTTCTTCTTCAACCACTCTTCAAGCGTCTTCTGAGAGCGCGTG 1440
QY 1704 CCGCGGACACCTGAGCTCAAGCGCAGCGTCTGCTGAGGAGCGCTTCAACATTTCCCGCAG 1763
Db 1441 CCGCGGACACCTGAGCTCAAGCGCAGCGTCTGCTGAGGAGCGCTTCAACATTTCCCGCAG 1500
QY 1764 CTTCCAGCTGCAAGTCCGCTCCACTGACCTTCACTCCACCGCAGACAGATGAAGAA 1823
Db 1501 CTTCCAGCTGCAAGTCCGCTCCACTGACCTTCACTCCACCGCAGACAGATGAAGAA 1560
QY 1824 GGCACCTTGAAGTGGTGGTGGCCAGGACCGTCTCCAGCGCTCAACAGTGGGATGGA 1883
Db 1561 GGCACCTTGAAGTGGTGGTGGCCAGGACCGTCTCCAGCGCTCAACAGTGGGATGGA 1620
QY 1884 CAGGGTAAATGCTCAGAGTGTACCTGAGGAGCGCAATTTACAGCGCTGCAATTTG 1943
Db 1621 CAGGGTAAATGCTCAGAGTGTACCTGAGGAGCGCAATTTACAGCGCTGCAATTTG 1680
QY 1944 TTTCCGAGTCTGCTCCCGGCCCCACGCTCACTTGTACTCATGCTGCTAAGATGCACAA 2003
Db 1681 TTTCCGAGTCTGCTCCCGGCCCCACGCTCACTTGTACTCATGCTGCTAAGATGCACAA 1740

QY 2004 CCGACACCCATACAACTACAGGGGCCACAAAGCAACTGCTGGGTAGCTTTCCACAG 2063
Db 1741 CCGACACCCATACAACTACAGGGGCCACAAAGCAACTGCTGGGTAGCTTTCCACAG 1800
QY 2064 ACATAAATATAGTCTCATCTGCAATCACAAGCACTAGCAGGTAAACCCACCACTCCCT 2123
Db 1801 ACATAAATATAGTCTCATCTGCAATCACAAGCACTAGCAGGTAAACCCACCACTCCCT 1860
QY 2124 GGATCTGAGGCCACACCGTGGGAGTCTGGCTGTACCTTCAAGGCCACAGAAACGGCA 2183
Db 1861 GGATCTGAGGCCACACCGTGGGAGTCTGGCTGTACCTTCAAGGCCACAGAAACGGCA 1920
QY 2184 CACATGTTCAAGTCTCACGCGCTCTCCATTCATTCGAATTTCTCAGTGTCCCTGCTCC 2243
Db 1921 CACATGTTCAAGTCTCACGCGCTCTCCATTCATTCGAATTTCTCAGTGTCCCTGCTCC 1980
QY 2244 GGTCTGGCACAGGACAGCATGCCCCCTCCGGGTCTGCGGCTACGCCACAGAGCTGTGCG 2303
Db 1981 GGTCTGGCACAGGACAGCATGCCCCCTCCGGGTCTGCGGCTACGCCACAGAGCTGTGCG 2040
QY 2304 TGTCTATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTACCCAGTACGATCGTCC 2363
Db 2041 TGTCTATGGCCCCCAACTCATGCTCCCTCTCTTGGCTACACCACTACCCAGTACGATCGTCC 2100
QY 2364 ACCGATGTCCACACACCCCCCAACCACTTGTCCACACAGTACCCAGTACGATCGTCC 2423
Db 2101 ACCGATGTCCACACACCCCCCAACCACTTGTCCACACAGTACCCAGTACGATCGTCC 2160
QY 2424 TGGCTCCCCAGAGTATCTTCCACTGAGACACGCGCCCCCAGAGGACAGTCCCCAG 2483
Db 2161 TGGCTCCCCAGAGTATCTTCCACTGAGACACGCGCCCCCAGAGGACAGTCCCCAG 2220
QY 2484 CCACCTCTGCAACTGACGCGCTCAGTCAACCCCTTTTAAAGCACCTCTGATTTACCAATG 2543
Db 2221 CCACCTCTGCAACTGACGCGCTCAGTCAACCCCTTTTAAAGCACCTCTGATTTACCAATG 2280
QY 2544 CAACACATCTGGTCTGCGATTTAGCAGAGATTTTGGACATACGAGGACCTCAGAC 2603
Db 2281 CAACACATCTGGTCTGCGATTTAGCAGAGATTTTGGACATACGAGGACCTCAGAC 2340
QY 2604 CGGAGGACACCTTCCCAACCCCAACACGCTGCTTATGTAACCGTGGAAAGCGGCCCT 2663
Db 2341 CGGAGGACACCTTCCCAACCCCAACACGCTGCTTATGTAACCGTGGAAAGCGGCCCT 2400
QY 2664 GCTGCCCTCCACACACATACATCACTCATCTACAGCCCTGTTCCGCGCTCAGAG 2723
Db 2401 GCTGCCCTCCACACACATACATCACTCATCTACAGCCCTGTTCCGCGCTCAGAG 2460
QY 2724 TCCCACTAGACCCAGTGGAGGGTTAGAGCAAGTAGGGGCCAGTTTCCCAATTCACC 2783
Db 2461 TCCCACTAGACCCAGTGGAGGGTTAGAGCAAGTAGGGGCCAGTTTCCCAATTCACC 2520
QY 2784 CTGTACGGAGTGGAGCGGATCTGAGTTCCTGCTGCTGCTTACGGTCCGGCTGGGAAT 2843
Db 2521 CTGTACGGAGTGGAGCGGATCTGAGTTCCTGCTGCTGCTTACGGTCCGGCTGGGAAT 2580
QY 2844 AAAGTTGTTTCTGGGCTTTAGCTTAAAAAATAAAAAA 2886
Db 2581 AAAGTTGTTTCTGGGCTTTAGCTTAAAAAATAAAAAA 2623

RESULT 13

US-10-127-900A-63

; Sequence 63, Application US/10127900A

; Publication No. US20030203429A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

Db	121	GGCCCGAGGACACCTGCCCCCGGCCCCACGCCGCTACCACTGCTGGGAAACCTCTCTGCA	180	1201	CCAGGGCAGGAGGTCTTTCCCTCTTGGCTCCATCTGTCATGACCCCAACATCTTCAA	1260
Qy	444	GCTACGGCCGGGCGCTGTATTACGGGCTCATGCGGCTAGTAAGAAAGTACGACCGGT	503	1524	GCACCCGAAGAGTTCAACCCAGACCGTTTCTTGATGACATGACGATGGATTCAGGAAGCA	1583
Db	181	GCTACGGCCGGGCGCTGTATTACGGGCTCATGCGGCTAGTAAGAAAGTACGACCGGT	240	1261	GCACCCGAAGAGTTCAACCCAGACCGTTTCTTGATGACATGACGATGGATTCAGGAAGCA	1320
Qy	504	GTTACCAATACCTTGGGACCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	563	1584	TGAGCGGTTCCTGCGCTTCTCTTAGGGAAGGTGTCTGCTTGGAGAGGCGCTGGCAAA	1643
Db	241	GTTACCAATACCTTGGGACCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	300	1321	TGAGCGGTTCCTGCGCTTCTCTTAGGGAAGGTGTCTGCTTGGAGAGGCGCTGGCAAA	1380
Qy	564	GCGGAGGCCCTTGGAGGTCAAGGCTGAGGAGTTACGCGCCGGGAAACCGTAGCGATGCT	623	1644	AGCGAGGTCTTTCCTCTTCTTCAACCATCTTCAAGCCTTCTCCCTGGAGAGCCGCTG	1703
Db	301	GCGGAGGCCCTTGGAGGTCAAGGCTGAGGAGTTACGCGCCGGGAAACCGTAGCGATGCT	360	1381	AGCGAGGTCTTTCCTCTTCTTCAACCATCTTCAAGCCTTCTCCCTGGAGAGCCGCTG	1440
Qy	624	GGAAGGACTTTTGTATGGCCATGGGGTTTCTTCTTCCAAACGGGAGCGGTGGAGCAGCT	683	1704	CCCGCCGACACCTGAGCCTCAAGCCACCGTCAAGTGGGCTTTTCAACATTTCCCCAGC	1763
Db	361	GGAAGGACTTTTGTATGGCCATGGGGTTTCTTCTTCCAAACGGGAGCGGTGGAGCAGCT	420	1441	CCCGCCGACACCTGAGCCTCAAGCCACCGTCAAGTGGGCTTTTCAACATTTCCCCAGC	1500
Qy	684	GAGGAAGTTTACATGCTTGTCTGCGGACCTTGGGCACTGGGCAATGGGAAAGCGAGGAGGA	743	1764	CTTCCAGCTGCAAGTCCGCTCCACTGACCTTCACTCCACCGAGACCCAGATGAAGAA	1823
Db	421	GAGGAAGTTTACATGCTTGTCTGCGGACCTTGGGCACTGGGCAATGGGAAAGCGAGGGA	480	1501	CTTCCAGCTGCAAGTCCGCTCCACTGACCTTCACTCCACCGAGACCCAGATGAAGAA	1560
Qy	744	GCTGATCAGGCGGAGGCCGCTGTCTGTGTGAGACATTCAGGGGACAGAAAGGACGCC	803	1824	GGCAACTTGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1883
Db	481	GCTGATCAGGCGGAGGCCGCTGTCTGTGTGAGACATTCAGGGGACAGAAAGGACGCC	540	1561	GGCAACTTGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1620
Qy	804	ATTGATCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	863	1884	CAGGGTTAATGTCTCCAGAGTGTACATGACGAGGACCAATTTACACGCTTCGATTTGT	1943
Db	541	ATTGATCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	600	1621	CAGGGTTAATGTCTCCAGAGTGTACATGACGAGGACCAATTTACACGCTTCGATTTGT	1680
Qy	864	TGGCCTCCGCTTCTCTATGAGATTAAGAGTTTCCAGGCGGTGGTCCGGGCGAGCTGGTG	923	1944	TTTTCCGGAGTCTGTCCACGGCCCAACGCTCACTTGAATCTGCTGCTTGAAGTGCACAA	2003
Db	601	TGGCCTCCGCTTCTCTATGAGATTAAGAGTTTCCAGGCGGTGGTCCGGGCGAGCTGGTG	660	1681	TTTTCCGGAGTCTGTCCACGGCCCAACGCTCACTTGAATCTGCTGCTTGAAGTGCACAA	1740
Qy	924	TACCTGCTGGAGTCACTCCAGGGGGTACAGCTACAGATGTTCTCTGCTGCTCT	983	2004	CCGCAACCCCTACACACTCAAGGGCCCAAGCAACTGCTGGGTAGCTTTCACACAG	2063
Db	661	TACCTGCTGGAGTCACTCCAGGGGGTACAGCTACAGATGTTCTCTGCTGCTCT	720	1741	CCGCAACCCCTACACACTCAAGGGCCCAAGCAACTGCTGGGTAGCTTTCACACAG	1800
Qy	984	GCGGCCCTTCCAGGCCCCCACAAGCAGCTCTCTCCACACAGCTCAGACCTTGGCTCCCTT	1043	2064	ACATAAATAGTTCATCTGCAATCAAGACATAGCCAGGTAAACCCACCACTCCCT	2123
Db	721	GCGGCCCTTCCAGGCCCCCACAAGCAGCTCTCTCCACACAGCTCAGACCTTGGCTCCCTT	780	1801	ACATAAATAGTTCATCTGCAATCAAGACATAGCCAGGTAAACCCACCACTCCCT	1860
Qy	1044	CACAGTCCGCGAGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1103	2124	GGATCTGACGCCCAACGCTGGGAGTCTGGCTGTCACTTCAACAGCCACAGAAACGGCCA	2183
Db	781	CACAGTCCGCGAGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	840	1861	GGATCTGACGCCCAACGCTGGGAGTCTGGCTGTCACTTCAACAGCCACAGAAACGGCCA	1920
Qy	1104	CCTTGTGATGCTTCTCTGCTGAAGATGGCAGAGGAAACAAACCCAGGCACAGAAAT	1163	2184	CACATGTTACAGCTCAGCGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT	2243
Db	841	CCTTGTGATGCTTCTCTGCTGAAGATGGCAGAGGAAACAAACCCAGGCACAGAAAT	900	1921	CACATGTTACAGCTCAGCGCCCTCTCCATTCATCGAACTTCTCAGTGTCCCTGTCCCT	1980
Qy	1164	CACCAACAGAAATGCTGATGACAGTCAATTTATTTGCTGCTGGGACGATGACGGT	1223	2244	GGTCCCTGGCAGAGGAAACAGCATGCCCTCCGGGTCTATGCCACCCAGAGACTGTGCG	2303
Db	901	CACCAACAGAAATGCTGATGACAGTCAATTTATTTGCTGCTGGGACGATGACGGT	960	1981	GGTCCCTGGCAGAGGAAACAGCATGCCCTCCGGGTCTATGCCACCCAGAGACTGTGCG	2040
Qy	1224	CAGCAACACGGTGGGTATACCTCTGCTCTGATGAATAACCTCATGTCCTCAAGATG	1283	2304	TGCTTATGGCCCCCAACTCATGCTCCCTCTCTGGCTACACCACTCTCCAGCCCTGTGACC	2363
Db	961	CAGCAACACGGTGGGTATACCTCTGCTCTGATGAATAACCTCATGTCCTCAAGATG	1020	2041	TGCTTATGGCCCCCAACTCATGCTCCCTCTCTGGCTACACCACTCTCCAGCCCTGTGACC	2100
Qy	1284	GSTACGTGAGAGCTGAATCGGAGCTGGGGCTGGCCAGGCACCAAGCCTTAGGGGACCG	1343	2364	ACCGATGTCCACACACCCCAACACTTGTGTCACACAGCTACCCAGCTACGATCGTCC	2423
Db	1021	GSTACGTGAGAGCTGAATCGGAGCTGGGGCTGGCCAGGCACCAAGCCTTAGGGGACCG	1080	2101	ACCGATGTCCACACACCCCAACACTTGTGTCACACAGCTACCCAGCTACGATCGTCC	2160
Qy	1344	TACCGGCTCCCTTACACGACGCGTCTGATGAGGCGAGCGGCTGTGGGCTGGT	1403	2424	TGGCTCCCGAGAGTATCTTCCCACTGAGACACGCCGCCCCCAACAGAGGCAAGTCCCGAG	2483
Db	1081	TACCGGCTCCCTTACACGACGCGTCTGATGAGGCGAGCGGCTGTGGGCTGGT	1140	2161	TGGCTCCCGAGAGTATCTTCCCACTGAGACACGCCGCCCCCAACAGAGGCAAGTCCCGAG	2220
Qy	1404	GCCCATGGGAATACCCGACCCCTCATGCGACCAACCGCTTCCGAGGCTACACCTGCC	1463	2484	CCACCTCTGCAAATGAGCGCTCAGTCAACCCCTTTTAAAGACCTTATCTACCAATG	2543
Db	1141	GCCCATGGGAATACCCGACCCCTCATGCGACCAACCGCTTCCGAGGCTACACCTGCC	1200	2221	CCACCTCTGCAAATGAGCGCTCAGTCAACCCCTTTTAAAGACCTTATCTACCAATG	2280
Qy	1464	CCAGGCGACGAGTCTTCCCTCTTGGCTCCATCTGCTGATGACCCCAACATCTTCAA	1523	2544	CAAAACATCTGGTCTGCGATTATGACAGAGACTTTGGACATACAGAGGACCTTCAGAC	2603
				2281	CAAAACATCTGGGCTCTGCGATTATGACAGAGACTTTGGACATACAGAGGACCTTCAGAC	2340

Db 668 AGCTCTCCATGGGGCCAGCCCTACGAGATGTTCTCTGGCTACTGTCAGCCCTGCGCAGGC 727
QY 1000 CCCACAGAGCTCTCCACAGCTAGCAGCTTGGCTGCTTCAAGTCCGGAGGTG 1059
Db 728 CCCACACAGCTCCAGCACCCTTGGGCGCCCTGGCTGCTTCACTATCCAGCAGGTA 787
QY 1060 CAGCAGCAGGAGGAACTGATGCTTTCGGGGCCCGCAGCTGACCTTGTGATGCTTTC 1119
Db 788 CAGAAACACAGGAGGACCTTCCAAACCTCAGTCTCGCAGTGATGCTTTCAGCCTTC 847
QY 1120 CTGCTGAGATGACAGAGGAGAAACAAACCCAGGACAGAAATTCACCAACAGAAACATG 1179
Db 848 CTGCTAAGATGACAGAGGAGAAACAAAGCCAGGATCAGAAATTCACCGAGAAGACTTG 907
QY 1180 CTGATGACAGTCAATTTATTTGCTGTTTGGGAGGATGACGCTGACGACCCAGCTGGC 1239
Db 908 CTGATGACGCTCATACTGCTGTTTGGGACCATGACCATCTGGTGGCCACATCCGC 967
QY 1240 TATACCTCTGCTCTGATGAATACCTCATGTCCAAAGTGGGTACGTGAGGAGCTG 1299
Db 968 TATGCCCTCTGCTCTGATGATACCTCAAGTCCAGCAGCGCTCCGGGAGGAGCTC 1027
QY 1300 AATCGGAGCTGGGCTGCGGAGGACCAAGACCTAGGAGACCTACCCGCTCCCTTAC 1359
Db 1028 ATACGAGAGCTGGGTCTGCGAGGGCTCCAACTCTCAGCGATCGAGTTCGCTCCCTTAC 1087
QY 1360 ACCGACGCTTCTGATGAGGCGCAGCGCTGCTGGCGCTGCTGCGATCGGATACCC 1419
Db 1088 ACGATGCGCTTTACAGAGGACACGCGCTCTTGGCATCTGGTACCATGGGATGCC 1147
QY 1420 CGACCTCATGCGGACCAACCGCTTCCGAGGGTACACCTGCCCCAGGCGACGAGGTC 1479
Db 1148 CACACCATCAGGAGGACCACTTGTCTCCGAGGGTACACTCTGCCAGGCGACTGAGTC 1207
QY 1480 TTCCCTCTGCTGCTCTGATGATGAGGACCAACATCTTCAAGCAGCAGGAGGTTTC 1539
Db 1208 TTCCCTCTGATGGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1267
QY 1540 AACCCAGACGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1599
Db 1268 CATCCAGGCGCTTCTGACGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1327
QY 1600 TTCTCTTAGGAGGAGCTGCTGCTTGGAGAGGCTTGGCAAGCGAGCTTCTTCTCTC 1659
Db 1328 TACTCTTAGTAAAGGAGTGTGCTTGGAGAGGCTTGGCTGGGAGGATGATGATGATGAT 1387
QY 1660 TTCTTACCAACATCTTCAAGCTTCTGCTGAGAGGCTGCTGCTGAGGAGGCTGCTGCT 1719
Db 1388 TTCTTCACTTCTTCAAGCTTCTGCTGAGAGGCTTCTGCTGAGAGGCTTCTGCTGAG 1447
QY 1720 AGCTTCAAGCCACCTCAGTGGCTTTCACATTTCCCGGCTTCCAGCTTCCAGCTGGA 1779
Db 1448 AGCTTCAAGCCACCTCAGTGGCTTTCACATTTCCCGGCTTCCAGCTTCCAGCTGGA 1507
QY 1780 CTGCTTCAAGCTTCACTTCACTTCA 1801
Db 1508 TGGCCCACTGGGACAGGATCCA 1529

RESULT 2
AK087069
LOCUS
DEFINITION
Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched library, clone:E030025C01 product:similar to CYTOCHROME P450 2S1 (Homo sapiens), full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK087069
GI:26352314
HTC; CAP trapper.
Mus musculus (house mouse)

REFERENCE
1
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

FEATURES
source

Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Itoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2292)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers
1. .2292
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:E030025C01"
/db_xref="MGI:2424609"

AUTHORS NTH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg. Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LECM2774 row: e column: 06
High quality sequence stop: 800.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

location/voadmills
1. 900
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6576174"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/notes="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."
```

ORIGIN

Query Match	30.1%;	Score 884.8;	DB 13;	Length 900;
Best Local Similarity	99.7%;	Pred. No. 3.8e-102;		
Matches 897;	Conservative	0;	Mismatches 2;	Indels 11;
				Gaps 1;

Qy	645	TTGGGGTTTCTTCTTCACACGGGGAGCGGTGGAGCGAGCTGAGGAAGTTTACCATGCTTGC	704
Db	1	TTGGGGTTTCTTCTTCC-ACGGGAGCGGTGGAGGCAGCTGAGGAAGTTTACCATGCTTGC	59

QY 705 TCTGCGGACCTGGGCATGGGAAGCGAAGCGCGAGAGCTGATCAGGCGGAGGCCG 764
Db 60 TCTGCGGACCTGGGCATGGGAAGCGAAGCGCGAGAGCTGATCAGGCGGAGGCCG 119

Qy	765	120	Db	824
	GTGTCGTGGAGACATTCAGGGGACAGAAAGACGCCCATTCGATCCCTCCCTCGTGCT	GTGTCGTGGAGACATTCAGGGGACAGAGAGCGCCCATTCGATCCCTCGTGCT	179	

Oy	825	GGCCAGGGCACCCTCCAAAGTAGTCGCTCCTCTTTTGSCCTCGCTTCTCCTATGA	884
Ddb	180	GGCCAGGGCACCCTCCAAAGTAGTCGCTCCTCTTTTGSCCTCGCTTCTCCTATGA	239

885	GGATAAGGAGTTC	CGAGCCGTGGTCCGGGCAGCTGGTGGTACCCCTGCTGGAGTCACTC	944
240	GGATAAGGAGTTC	CGAGCCGTGGTCCGGGCAGCTGGTGGTACCCCTGCTGGAGTCACTC	299

QY	db	945	300	1,004
		CCAGGGGGT	CAGACTACAGATGTTCTCTGTGTTCTCTGCGCCCTGCCAGGCCCA	1,004
			CAGGGGGT	359

QY	1005	360	db
	CAAGCAGCTCCTCCACCAAGTCAGCACCCTTGCTGCTTCACAGTCGGCAGGTGCAGCA	CAAGCAGCTCCTCCACCAAGTCAGCACCCTTGCTGCTTCACAGTCGGCAGGTGCAGCA	1064
			419

QY	1065	GCACAGGGGAACCTGGATGCTTCGGGCCCCGACCTGTGATGCTTCTCTGCT	1124
Db	420	GCACAGGGGAACCTGGATGCTTCGGGCCCCGACCTGTGATGCTTCTCTGCT	479

1125	GAAGATGGCAGCAGAGGACAAAACCCAGGCACAGAAATTCACCAACAGAA	CATGCTGAT	1184
QY			
480	GAAGATGGCAGCAGAGGACAAAACCCAGGCACAGAAATTCACCAACAGAA	CATGCTGAT	539
Db			

QY	1185	GACAGTCATTTATTTGCTGTTTGTGGAGCATGACGGTCAGCACCAACGGTCGGCTATAC	1244
Db	540	GACAGTCATTTATTTGCTGTTTGTGGAGCATGACGGTCAGCACCAACGGTCGGCTATAC	599
QY	1245	CCTCCTGCTCCTGATGAAATACCTCATGTCTCCAAAAGTGGTACGTGAGGAGCTGAATCG	1304
Db	600	CCTCCTGCTCCTGATGAAATACCTCATGTCTCCAAAAGTGGTACGTGAGGAGCTGAATCG	659
QY	1305	GGAGCTGGGGCTGGCCAGGCACCAAGCCTAGGGGACCGGTACCCGCCCTCCCTTACACCGA	1364
Db	560	GGAGCTGGGGCTGGCCAGGCACCAAGCCTAGGGGACCGGTACCCGCCCTCCCTTACACCGA	719
QY	1365	CGGGTTTCTCATGAGCGCAGCGGCTGCTGGGGCTGGTCCCATGGGAATACCCCGCAC	1424
Db	720	CGGGTTTCTCATGAGCGCAGCGGCTGCTGGGGCTGGTCCCATGGGAATACCCCGCAC	779
QY	1425	CCTCATGGGACACACCGCTTCCGAGGGTACACCTGTCCCCAGGGCACGAGGTCCTTCCC	1484
Db	780	CCTCATGGGACACACCGCTTCCGAGGGTACACCTGTCCCCAGGGCACGAGGTCCTTCCC	839
QY	1485	CCTCCTTGGCTCCTCCTGATGACCCCAACATCTTCAAGCACCCAGAGAGTTCACCC	1544
Db	840	CCTCCTTGGCTCCTCCTGATGACCCCAACATCTTTCAGACCCAGAGAGTTCACCC	899
RESULT	4		
BO928732			
LOCUS		884 bp	linear
DEFINITION	AGENCOCURT_10037033 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6480653	mRNA	EST 20-AUG-2002
	5', mRNA sequence.		

ACCESSION	BQ928732
VERSION	BQ928732.1
KEYWORDS	GI:22343763
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 884)
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2660 row: m column: 06
High quality sequence stop: 792.

FEATURES	SOURCE
1. Age	1. Age
2. Gender	2. Gender
3. Marital Status	3. Marital Status
4. Education	4. Education
5. Income	5. Income
6. Occupation	6. Occupation
7. Religion	7. Religion
8. Political Affiliation	8. Political Affiliation
9. Health Status	9. Health Status
10. Travel History	10. Travel History
11. Employment Status	11. Employment Status
12. Family Size	12. Family Size
13. Home Ownership	13. Home Ownership
14. Vehicle Ownership	14. Vehicle Ownership
15. Insurance Status	15. Insurance Status
16. Substance Use	16. Substance Use
17. Mental Health	17. Mental Health
18. Chronic Conditions	18. Chronic Conditions
19. Vaccination Status	19. Vaccination Status
20. Comorbidities	20. Comorbidities

```

location/Qualifiers
1. .884
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6480653"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site.1: XhoI;
Site.2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using
the following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."

```

Query Match 29.6%; Score 871; DB 13; Length 884;
Best Local Similarity 99.8%; Pred. No. 2.1e-100;

Matches 882; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	
Qy	1323 GGCACCAAGCCTAGGGACCGTACCGGCTCCCTTACACGACGGGTTCGATGAGC 1382
Db	1 GGCACCAAGCCTAGGGACCGTACCGGCTCCCTTACACGACGGGTTCGATGAGC 60
Qy	1383 GCAGGGCTGTGGCGTGTGGCCATCGGGAATACCCGACCCCTCATGCGGACACCG 1442
Db	61 GCAGGGCTGTGGCGTGTGGCCATCGGGAATACCCGACCCCTCATGCGGACACCG 120
Qy	1443 CTTCCGAGGTACACCTGCCCGCCAGGACGAGGTCTTCCCTCTTGGCTTCATCCT 1502
Db	121 CTTCCGAGGTACACCTGCCCGCCAGGACGAGGTCTTCCCTCTTGGCTTCATCCT 180
Qy	1503 GCATGACCCCAACATCTTCAAGCACCCAGAGAGGTTCACCCAGACCGTTCCTGGATGC 1562
Db	181 GCATGACCCCAACATCTTCAAGCACCCAGAGAGGTTCACCCAGACCGTTCCTGGATGC 240
Qy	1563 AGATGGACGTTTCAGGAAGCATGAGCGTCTCTGCCCTTCTCTTAGGGAAGCGTCTG 1622
Db	241 AGATGGACGTTTCAGGAAGCATGAGCGTCTCTGCCCTTCTCTTAGGGAAGCGTCTG 300
Qy	1623 CTTTGGAGGGCTGGCAAAAGCGAGCTCTTCTCTTCTTCCACCACTCTACAAGC 1682
Db	301 CTTTGGAGGGCTGGCAAAAGCGAGCTCTTCTCTTCTTCCACCACTCTACAAGC 360
Qy	1683 CTTTCTCTGGAGACCGGTGCCCGGACACCTGAGCTTCAAGCCACCGTCAGTGG 1742
Db	361 CTTTCTCTGGAGACCGGTGCCCGGACACCTGAGCTTCAAGCCACCGTCAGTGG 420
Qy	1743 CTTTTCACATCTCCCGACCTTCCAGTGAAGTCCGTCCTCCACTGACCTTCACTCCAC 1802
Db	421 CTTTTCACATCTCCCGACCTTCCAGTGAAGTCCGTCCTCCACTGACCTTCACTCCAC 480
Qy	1803 CACGACACACGATGAAGGAAGCAACTTGGAGTGTGGGTGCCGAGGACGGTGCCTCC 1862
Db	481 CACGACACACGATGAAGGAAGCAACTTGGAGTGTGGGTGCCGAGGACGGTGCCTCC 540
Qy	1863 AGCCTCAACAGTGGGATGGAAGGTTAATGTCTCCAGAGTGTACCTGAGGAGCA 1922
Db	541 AGCCTCAACAGTGGGATGGAAGGTTAATGTCTCCAGAGTGTACCTGAGGAGCA 600
Qy	1923 CATTTACACCTGACGTTGTTTCCGGAGTCTGCCACGCGCCACACGCTCACTTGAC 1982
Db	601 CATTTACACCTGACGTTGTTTCCGGAGTCTGCCACGCGCCACACGCTCACTTGAC 660
Qy	1983 TCATGCTGCTAAGATGCACACCGCACACCCATACACAACTACAAGGGCCACAAGCAAC 2042
Db	661 TCATGCTGCTAAGATGCACACCGCACACCCATACACAACTACAAGGGCCACAAGCAAC 720
Qy	2043 TGTGGGTTAGTTTCCACAGACATAAATATAGTCATCTGAAATCACAAGCACATAGCC 2102
Db	721 TGTGGGTTAGTTTCCACAGACATAAATATAGTCATCTGAAATCACAAGCACATAGCC 780
Qy	2103 AGGTAAACCAACCTCCCTGATCTGCAGCCACACCTT-GGGAGTCTGGGTGTCACT 2161
Db	781 AGGTAAACCAACCTCCCTGATCTGCAGCCACACCTGAGGAGTCTGGGTGTCACT 840
Qy	2162 TCACAGCCACAGAAACGGCCACACATGTTTCAGCTCACAGC 2205
Db	841 TCACAGCCACAGAAACGGCCACACATGTTTCAGCTCACAGC 884

RESULT 5
BQ895833
LOCUS BQ895833
DEFINITION AGENCOURT_8675453 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380072
5', mRNA sequence.
ACCESSION BQ895833
VERSION BQ895833.1 GI:22287847
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 899)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LUCM2568 row: b column: 09
High quality sequence stop: 644.
Location/Qualifiers
1. 899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380072"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN

Query Match 29.4%; Score 867; DB 13; Length 899;	
Best Local Similarity 99.1%; Pred. No. 6.6e-100;	
Matches 892; Conservative 0; Mismatches 6; Indels 2; Gaps 2;	
Qy	1323 GGCACCAAGCCTAGGGACCGTACCGGCTCCCTTACACGACGGGTTCGATGAGC 1382
Db	1 GGCACCAAGCCTAGGGACCGTACCGGCTCCCTTACACGACGGGTTCGATGAGC 59
Qy	1383 GCAGCGGCTGTGGCGTGTGGCCATCGGGAATACCCGACCCCTCATGCGGACACCGC 1442
Db	60 GCAGCGGCTGTGGCGTGTGGCCATCGGGAATACCCGACCCCTCATGCGGACACCGC 119
Qy	1443 CTTCCGAGGTTACACCTTGCCCCAGGCGCACGGAGGTCTTCCCTCTCTTGGTCCATCT 1502
Db	120 CTTCCGAGGTTACACCTTGCCCCAGGCGCACGGAGGTCTTCCCTCTCTTGGTCCATCT 179
Qy	1503 GCATGACCCCAACATCTTCAAGCACCCAGAGGTTCACCCAGACCGTTCCTGATGC 1562
Db	180 GCATGACCCCAACATCTTCAAGCACCCAGAGGTTCACCCAGACCGTTCCTGATGC 239
Qy	1563 AGATGACGTTTCAGGAAGCATGAGCGTCTCTGCCCTTCTCTTAGGGAAGCGTCTCTG 1622
Db	240 AGATGACGTTTCAGGAAGCATGAGCGTCTCTGCCCTTCTCTTAGGGAAGCGTCTCTG 299
Qy	1623 CTTTGGAGGGCTGGCAAAAGCGGAGCTCTTCTCTTCAACCACTCTCAAGC 1682
Db	300 CTTTGGAGGGCTGGCAAAAGCGGAGCTCTTCTCTTCTTCAACCACTCTCAAGC 359
Qy	1683 CTTTCTCTGGAGAGCGCTGCCCGCGACACCTTGAGCTCAAGCCACCGTCAGTGG 1742
Db	360 CTTTCTCTGGAGAGCGCTGCCCGCGCGACACCTTGAGCTCAAGCCACCGTCAGTGG 419
Qy	1743 CTTTTTCAACATTTCCCGCCAGCCTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCAC 1802
Db	420 CTTTTTCAACATTTCCCGCCAGCCTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCAC 479
Qy	1803 CACGACACACGATGAAGGAAGCAACTTGGAGTGTGGGTGCCGAGGACGGTGCCTCC 1862

Db 480 CACGACACAGATGAAGGAAGCAACTTGGAGTGGTGGTCCCGAGGAGGTCCTCC 539
QY 1863 AGCCTCAACAGTGGGCAATGACAGGGTTAATGTCTCAGAGTGTACACTGAGGAGCCCA 1922
Db 540 AGCCTCAACAGTGGGCAATGACAGGGTTAATGTCTCAGAGTGTACACTGAGGAGCCCA 599
QY 1923 CATTTACACGCTTGCAGTCTGTTTCCGGAGTCTGTCCACGGCCACACGCTCAGTTGAC 1982
Db 600 CATTTACACGCTTGCAGTCTGTTTCCGGAGTCTGTCCACGGNCCACACGCTCAGTTGAC 659
QY 1983 TCATGCTGTAAAGATGACAAACCGCACACCCATACACAACTACAAGGGCCACAAAAGCAAC 2042
Db 660 TCATGCTGTAAAGATGACAAACCGCACACCCATACACAACTACAAGGGCCACAAAAGCAAC 719
QY 2043 TGTGGGTAGCTTTCCACAGACATAAATATAGTGCATCTGCAATCACAAGCAGCATAGCC 2102
Db 720 TGTGGGTAGCTTTCCACAGACATAAATATAGTGCATCTGCAATCACAAGCAGCATAGCC 779
QY 2103 AGGTAAACCCACCAACTCCCTGATCTGCAGCCACACGTTGGAGTCTGGCTGCACCTT 2162
Db 780 AGGTAAACCCACCAACTCCCTGATCTGCAGCCACACGTTGGAGTCTGGCTGCACCTT 839
QY 2163 CACAAGCCACAGAAACGGCCACACATGTTTCAAGTTCARAGCCCTCTCCATTCATCGA 2221
Db 840 CACAAGCCACAGAAACGGCCACACATGTTTCAAGTTCARAGCCCTCTCCATTCATCGA 899

RESULT 6

BQ920829
LOCUS
DEFINITION BQ920829 945 bp mRNA linear EST 20-AUG-2002
AGENCOURT 10033015 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6482552
5', mRNA sequence.

ACCESSION BQ920829.1 GI:22335673
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 945)
NIH-MGC <http://imgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LCM2665 row: 1 column: 09

High quality sequence stop: 643.

Location/Qualifiers

1..945

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6482552"

/tissue_type="carcinoma, cell line"

/lab_host="PH10B (phage-resistant)"

/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

BU849902 924 bp mRNA linear EST 16-OCT-2002

RESULT 7

BU849902

LOCUS

Query Match 29.4%; Score 865.2; DB 13; Length 945;
Best Local Similarity 97.5%; Pred. No. 1.1e-99;
Matches 909; Conservative 0; Mismatches 20; Indels 3; Gaps 3;
QY 1328 CAAGCCTAGGGGACCGTACCGCGCTCCCTTACACCGACGCGGTTCTGCATGAGCGGAGC 1387
Db 5 CGAGCCTAGGGGACCGTACCGCGCTCCCTTACACCGACGCGGTTCTGCATGAGCGGAGC 64
QY 1388 GGCTGCTGGCGCTGGTCCCATGGGAATACCCCGCACCTCATCGGACACACCGCTTTC 1447
Db 65 GGCTGCTGGCGCTGGTCCCATGGGAATACCCCGCACCTCATCGGACACACCGCTTTC 124
QY 1448 GAGGGTACACCTTGCCTCCAGGGACCGAGGTCCTCCCTTCTTGGTCCCATCTGCAATG 1507
Db 125 GAGGGTACACCTTGCCTCCAGGGACCGAGGTCCTCCCTTCTTGGTCCCATCTGCAATG 184
QY 1508 ACCCCACACATCTTCAAGCAACCCAGAGAGTTCAACCCAGACCGGTTTCTTGGATGCAATG 1567
Db 185 ACCCCACACATCTTCAAGCAACCCAGAGAGTTCAACCCAGACCGGTTTCTTGGATGCAATG 244
QY 1568 GACGGTTCAGGAAGCATGAGGCGTTCCTGCGCTTCTCTTAGGGAAGCGTGTCTGCGCTT 1627
Db 245 GACGGTTCAGGAAGCATGAGGCGTTCCTGCGCTTCTCTTAGGGAAGCGTGTCTGCGCTT 304
QY 1628 GAGAGGCGCTGGCAAAAGCGGAGCTCTTCTCTTCTTACCAATCTTAAGCGCTTCT 1687
Db 305 GAGAGGCGCTGGCAAAAGCGGAGCTCTTCTCTTCTTACCAATCTTAAGCGCTTCT 364
QY 1688 CCTGTGAGAGCGCGTCCCGCGGACACCTGAGCCTCAAGCCACCGTTCAGTGGCGCTT 1747
Db 365 CCTGTGAGAGCGCGTCCCGCGGACACCTGAGCCTCAAGCCACCGTTCAGTGGCGCTT 424
QY 1748 TCAACATTTCCCGCAGCTTCCAGCTGCAAGTCCGTCCTTCCCTTACCTTCACTCCACACGC 1807
Db 425 TCAACATTTCCCGCAGCTTCCAGCTGCAAGTCCGTCCTTCCCTTACCTTCACTCCACACGC 484
QY 1808 AGACAGATGAGGAAGGCAACTTGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1867
Db 485 AGACAGATGAGGAAGGCAACTTGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 544
QY 1868 CAACAGTGGGATGAGGAGGTTAATGTCTCCAGAGTGTACACTGAGGAGGAGGAGGAGGAG 1927
Db 545 CAACAGTGGGATGAGGAGGTTAATGTCTCCAGAGTGTACACTGAGGAGGAGGAGGAGGAG 604
QY 1928 ACAAGCCTGCAAGTGTGTTTCCGAGTGTCTTCCACCGGCGCCACACGCTTCACTTGA 1987
Db 605 ACAAGCCTGCAAGTGTGTTTCCGAGTGTCTTCCACCGGCGCCACACGCTTCACTTGA 664
QY 1988 CTGCTAAGATGACAAACCGCCACACCCATACAACTACAAGGGGCCACAAAGCAATGCTG 2047
Db 665 CTGCTAAGATGACAAACCGCCACACCCATACAACTACAAGGGGCCACAAAGCAATGCTG 724
QY 2048 GGTAGCTTTCCACAGACATTAATAGTCCATCTGCAATCACAAGCAGCATAGCCAGGTA 2107
Db 725 GGTAGCTTTCCACAGACATTAATAGTCCATCTGCAATCACAAGCAGCATAGCCAGGTA 784
QY 2108 ACCCACCAACT-CCGCTGGATCTGCAGGCCACACGTTGGAGTCTGGCTCTCACTTTCACA 2166
Db 785 ACCCACCAACTCCCCCTGGATCTGCAGGCCACACGTTGGAGTCTGGCTCTCACTTTCACC 844
QY 2167 ACCCACAGAAACGGCCACACATGTTTACA-GTCAAGCGCCTCTCCATTCAATCG-AACT 2224
Db 845 ACCCACAGAAACGGCCACACATGTTTACAAGCGGCTTCAAGCGGCTTCACTTATCGAACT 904
QY 2225 TCTCAGTGTCCCTGTCCCTGGTGGCTGGCACA 2256
Db 905 TCTCAGTGTCCCTGTCCCTGGTGGCTGGCACA 936

```
DEFINITION  AGENCOURT 10441031 NIH_MGC 109 Homo sapiens cDNA clone
IMAGE:6598606 5', mRNA sequence.
ACCESSION    BU849902.1 GI:24034865
VERSION      BU849902.1 GI:24034865
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 924)
              NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE        Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM2831 row: k column: 22
              High quality sequence stop: 701.
              Location/Qualifiers
                1..924
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6598606"
                  /tissue_type="teratocarcinoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_109"
                  /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
                  XhoI; cDNA made by oligo-dT priming. Directionally cloned
                  into EcoRI/XhoI sites using the following 5' adaptor:
                  GGCAGGAG(G). Library constructed by Ling Hong in the
                  laboratory of Gerald M. Rubin (University of California,
                  Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                  Superscript II RT (Life Technologies). Note: this is a
                  NIH_MGC Library."
ORIGIN
  Query Match      29.3%; Score 861.6; DB 13; Length 924;
  Best Local Similarity 98.6%; Pred No. 3,1e-99;
  Matches 900; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
QY 290 GCGACCTGCCGAGATGAGGCGACCGGCACCTGGCGCGTCTGCTGGCGTGGCGCTGC 349
DB 1 GCGGACCTGCCGAGATGAGGCGACCGGCACCTGGCGCGTCTGCTGGCGTGGCGCTGC 60
QY 350 TCCTGCTCTACGCTGGCGTCTGTCGGGACCGAGGCCCGAGGCCACCTGCCCGCGGGC 409
DB 61 TCCTGCTCTACGCTGGCGTCTGTCGGGACCGAGGCCCGAGGCCACCTGCCCGCGGGC 120
QY 410 CCAGCCGCTACCACTGCTGGGAAACCTCTCGAGCTACGGCCCGGGCGGTGATTGAG 469
DB 121 CCAGCCGCTACCACTGCTGGGAAACCTCTCGAGCTACGGCCCGGGCGGTGATTGAG 180
QY 470 GGCTCATGCGGTGAGTAAAGTAGTACGACCGGTGTTCACCATCTACCTGGGACCTGGC 529
DB 181 GGCTCATGCGGTGAGTAAAGTAGTACGACCGGTGTTCACCATCTACCTGGGACCTGGC 240
QY 530 GGCCTGTGGTGTCTGCTGGGACGAGGCTGTGGGAGCGCCCTGGAGGTCAAGGTG 589
DB 241 GGCCTGTGGTGTCTGCTGGGACGAGGCTGTGGGAGCGCCCTGGAGGTCAAGGTG 300
QY 590 AGGAGTTACGGCGCGGGAACCGTAGCATGCTCGAAGCGGACTTTTGATGCCATGGGG 649
DB 301 AGGAGTTACGGCGCGGGAACCGTAGCATGCTCGAAGCGGACTTTTGATGCCATGGGG 360
QY 650 TTTTCTTCTCCAAACGGGAGCGGTGGAGGACGCTGAGGAAAGTTTACCATGCTTGTCTGC 709
|||||
```

```
Db 361 TTTTCTTCTCCAAACGGGAGCGGTGGAGGACGCTGAGGAAAGTTTACCATGCTTGTCTGC 420
QY 710 GGGACCTGGGCATGGGGAAGCGAGAAGCGGAGGAGCTGATCCAGCGGAGGCCCGGTGC 769
Db 421 GGGACCTGGGCATGGGGAAGCGAGAAGCGGAGGAGCTGATCCAGCGGAGGCCCGGTGC 480
QY 770 TGGTGGAGACATTCCAGGGGACAGAAGGACGCCCATTCGATCCCTCCCTGCTGTGGGCC 829
Db 481 TGGTGGAGACATTCCAGGGGACAGAAGGACGCCCATTCGATCCCTCCCTGCTGTGGGCC 540
QY 830 AGGCACCTCCAAAGTAGTCTGCTCCCTCTCTTTGGCCTCGGCTTCTCTATGAGGATA 889
Db 541 AGGCACCTCCAAAGTAGTCTGCTCCCTCTCTTTGGCCTCGGCTTCTCTATGAGGATA 600
QY 890 AGGAGTTCCAGGCGGTGTCGGGCGAGCTGGTGTTACCTCTGCGGAGTCACTGCCAGG 949
Db 601 AGGAGTTCCAGGCGGTGTCGGGCGAGCTGGTGTTACCTCTGCGGAGTCACTGCCAGG 660
QY 950 GGGGTTCAGACCTACGAGATGTTCTCTGTTCTGCGGCCCTCGCAGGCCGCCCAAGC 1009
Db 661 GGGGTTCAGACCTACGAGATGTTCTCTGTTCTGCGGCCCTCGCAGGCCGCCCAAGC 720
QY 1010 AGCTCTCCACACGCTACGACCTTGGTGGCTTTCACAGTCCGGGAGTGGAGGACACC 1069
Db 721 AGCTCTCCACACGCTACGACCTTGGTGGCTTTCACAGTCCGGGAGTGGAGGACACC 780
QY 1070 AGGGG-AACCTGGATGCTTCGGGCCCGCAGCTGACCTTTCGATGCTT-CCTGCTGAA 1127
Db 781 AGGGGAAACCTGGATGCTTCGGGCCCGCAGCTGACCTTTCGATGCTT-CCTGCTGAA 840
QY 1128 GATGCGACAGGAGGAACAAACCCAGGCGACAGAA-TTCACCAACAAGAACATGCTGATGA 1186
Db 841 GAGGCGACCGGAGGACAAACCCAGGCGACAGAA-TTCACCAACAAGAACATGCTGATGA 900
QY 1187 CAGTCATTATT 1199
Db 901 CAGTCATTATT 913
RESULT 8
BQ900302
LOCUS       AGENCOURT_8675726 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380105
DEFINITION  5', mRNA sequence.
ACCESSION   BQ900302
VERSION     BQ900302.1 GI:22292316
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 903)
              NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: DCTD/DTF
              cDNA Library Preparation: Rubin Laboratory
              DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Place: LLCM2568 row: c column: 18
              High quality sequence stop: 690.
              Location/Qualifiers
                1..903
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6380105"
                  /tissue_type="carcinoma, cell line"
FEATURES
  source
```

```

/lab_host="DH10B (phage-resistant) "
/clone_lib="NIH MGC 40"
/notes="Organ: prostate; Vector: pORF7; Site.1: XhoI;
Site.2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using Zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC library."

```

ORIGIN

Query Match		28.8%	Score 848.4	DB 13	Length 903
Best Local Similarity		98.1%	Pred. No. 1.4e-97		
Matches	888	Conservative	0	Mismatches	14
				Indels	3
				Gaps	3
QY	645	TGGGGTTTCTTCTCCAA	CGGGAGCGGTGGAGGCAGCTGAGGAAGTTTACCATGCTTGC	704	
Db	1	TGGGGTTTCTTCTCC	ACGGGAGCGGTGGAGGCAGCTGAGGAA-TTTTACCATGCTTGC	58	
QY	705	TCTCGGGGACCTGGGCAT	TGGGAAGCAGAAAGCGGAGAGCTGATCCAGGCGGAGGCGCG	764	
Db	59	TCTCGGGACCTGGGCAT	TGGGAAGCAGAAAGCGGAGAGCTGATCCAGGCGGAGGCGCG	118	
QY	765	GTCTCTGTGGAGACAT	TCAGGGGACAGAAAGACGCCATTGATCCCTCCCTGCTGCT	824	
Db	119	GTCTCTGTGGAGACAT	TCCAGGGGACAGAAAGAGGCCOATTCGATCCCTCCCTGCTGCT	178	
QY	825	GGCCGAGGCCACTCCAA	CGTAGTCTGCTCCCTCTCTTTTGGCGCTCCGCTTCTCCTATGA	884	
Db	179	GGCCCGAGGCCACTCCAA	CGTAGTCTGCTCCCTCTCTTTTGGCGCTCCGCTTCTCCTATGA	238	
QY	885	GGATAAGGAGTTCAG	GCCGCTGTCGGGCAGCTGGTGTAACCTCGTCGGGCGCCCTG	1004	
Db	239	GGATAAGGAGTTCAG	GCCGCTGTCGGGCAGCTGGTGTAACCTCGTCGGGCGCCCTG	358	
QY	945	CGAGGGGGTCAGAC	CTACGAGATGTTCTCCTGGTTCTCTCGGCCCTCGCAGGCGCCCTG	1004	
Db	299	CGAGGGGGTCAGAC	CTACGAGATGTTCTCCTGGTTCTCTCGGCCCTCGCAGGCGCCCTG	358	
QY	1005	CAAGCAGCTCCTCCAC	CGGTACAGCACTTGGCTGCTTACAGTCCGGCAGGTGAGCA	1064	
Db	359	CAAGCAGCTCCTCCAC	CGGTACAGCACTTGGCTGCTTACAGTCCGGCAGGTGAGCA	418	
QY	1065	GCACCAAGGGGAACT	CGATGCTTTCGGGCCCGCAGCTGACCTTGTCGATGCTCCCTGCT	1124	
Db	419	GCACCAAGGGGAACT	CGATGCTTTCGGGCCCGCAGCTGACCTTGTCGATGCTCCCTGCT	478	
QY	1125	GAGATGGCACAGGAG	GAACAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGAT	1184	
Db	479	GAGATGGCACAGGAG	GAACAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGAT	538	
QY	1185	GACAGTCATTTATTT	GCTTGTCTGGGACGATGACGCTCAGCACACCGTCCGCTATAC	1244	
Db	539	GACAGTCATTTATTT	GCTTGTCTGGGACGATGACGCTCAGCACACCGTCCGCTATAC	598	
QY	1245	CCTCTGCTCCTGATGA	AAATACCTCATGTCCTCAAAAGTGGGTACGTGAGGAGCTGAATCG	1304	
Db	599	CCTCTGCTCCTGATGA	AAATACCTCATGTCCTCAAAAGTGGGTACGTGAGGAGCTGAATCG	658	
QY	1305	GGAGTGGGGGTGGCC	AGGACCAAGCTAGGGGACCGTACCCGCTCCCTTACACCGA	1364	
Db	659	GGAGTGGGGGTGGCC	AGGACCAAGCTAGGGGACCGTACCCGCTCCCTTACACCGA	718	
QY	1365	CGGGTTCTGCATGAG	CGCAGCGGTGCTGGCGCTGGTCCCATGGGAATACCCCGCAC	1424	
Db	719	CGGGTTCTGCATGAG	CGCANGCANGCGGTGCTGGCGCTGGTCCCATGGGAATACCCCGCAC	778	
QY	1425	CCTCATGCGGACCA	CCCGCTTCCGAGGGTACACCTTGCCCGCAGGGGACGAGGTCTT-CC	1483	
Db	779	CCTCATGCGGACCA	CCCGCTTCCGAGGGTACACCTTGCCCGCANGGACGAGGAGGTCTTCC	838	
QY	1484	CCCTCCTTGGCTCCA	TCCTGTCATGACCCCCAACATCTTCAAGCACCCGAGAAGAGTTCAAC	1543	

Db	839	CCCTCTTGGCTCCATCCTGATGACCCCCACATCTTCCAGACACCAGAGAGTTCAACC	898
QY	1544	CAGAC	1548
Db	899	CAGAC	903
RESULT 9			
BUS52425			
LOCUS			
DEFINITION	BUS52425	915 bp	mRNA linear EST 16-SEP-2000
ACCESSION	AGENCOURT 10333108	NIH_MGC_40	Homo sapiens cDNA clone IMAGE:6576204
VERSION	BUS52425		5', mRNA sequence.
KEYWORDS	BUS52425.1	GI:22902697	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 915)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: DCTD/DTF		
	cDNA Library Preparation: Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Plate: LLCM2774	row: f	column: 12
	High quality sequence stop: 754.		
FEATURES	Location/Qualifiers		
source	1..915		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6576204"		
	/tissue_type="carcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_40"		
	/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."		

ORIGIN

	Query Match	28.3%	Score 833;	DB 13;	Length 915;
	Best Local Similarity	99.0%;	Pred. No. 1.2e-95;		
	Matches 858;	Conservative 0;	Mismatches 7;	Indels 2;	Gaps 2;
QY	1323	GGCACCAAGCCTAGGGGACCGTACCGCGCTCCCTTACACGACGCGGTTCTGCATGAGGC	1382		
Db	1	GGCACCAAGCCTAGGGGACCGTACCGCGCTCCCTTACACGACGCGGTTCTGCATGAGGC	60		
QY	1383	GCAGCGCTGCTGGCGCTGGTGCCCATGGGAATACCCCGCACCCCTCATCGGAGCACCCCG	1442		
Db	61	GCAGCGCTGCTGGCGCTGGTGCCCATGGGAATACCCCGCACCCCTCATCGGAGCACCCCG	120		
QY	1443	CTTCCGAGGGTACACCCCTGC GCCCAGGCGACGGAGGTCTTCCCCCTCTCTTGCTCCCATCT	1502		
Db	121	CTTCCGAGGGTACACCCCTGCCCCAGGCGACGGAGGTCTTCCCCCTCTCTTGCTCCCATCT	180		
QY	1503	GCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGATGC	1562		
Db	181	GCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGATGC	240		

```
Qy 1563 AGATGGACGGTTTCAGGAAGCATGAGCGTTCCTGCGCTTCTCTTAGGGAAGCGTGTCTG 1622
Db 241 AGATGGACGGTTTCAGGAAGCATGAGCGTTCCTGCGCTTCTCTTAGGGAAGCGTGTCTG 300
Qy 1623 CTTTGGAGAGCGCTTGGCAAGAGCGAGCTCTTCTCTTCTTTCACCAATCTCTTACCAAGC 1682
Db 301 CTTTGGAGAGCGCTTGGCAAGAGCGAGCTCTTCTCTTCTTTCACCAATCTCTTACCAAGC 360
Qy 1683 CTTTCTCTTGGAGAGCGCTTGGCCCGGACACCTCTGAGCTTCAAGCCACCGTCAAGTGG 1742
Db 361 CTTTCTCTTGGAGAGCGCTTGGCCCGGACACCTCTGAGCTTCAAGCCACCGTCAAGTGG 420
Qy 1743 CTTTCTTCAACATCTCCACAGCTTCCAGTGCAGTCCGTCCTCCACTGACCTTCACTCCAC 1802
Db 421 CTTTCTTCAACATCTCCACAGCTTCCAGTGCAGTCCGTCCTCCACTGACCTTCACTCCAC 480
Qy 1803 CACGACAGACCATGAGGAAGGCAACTTGAAGTGGTGGTGCCCGGACGCGTGCCTCC 1862
Db 481 CACGACAGACCATGAGGAAGGCAACTTGAAGTGGTGGTGCCCGGACGCGTGCCTCC 540
Qy 1863 AGCCTCAACAGTGGCATGGACAGGGTTAATGTCTTCCAGAGTGTACATGCGAGGACGCA 1922
Db 541 AGCCTCAACAGTGGCATGGACAGGGTTAATGTCTTCCAGAGTGTACATGCGAGGACGCA 600
Qy 1923 CATTTACAGCGCTGCAGTGTGTTTCCGGAGTCTGTCCACGCGCCACACGCTCACTTGAC 1982
Db 601 CATTTACAGCGCTGCAGTGTGTTTCCGGAGTCTGTCCACGCGCCACACGCTCACTTGAC 660
Qy 1983 TCATGCTGTAAAGTGCACAAACGCGACACCCATACACAACTACAAGGGGCAACAAAGCAAC 2042
Db 661 TCATGCTGTAAAGTGCACAAACGCGACACCCATACACAACTACAAGGGGCAACAAAGCAAC 720
Qy 2043 TGCTGGGTAGCTTTCACAGACATAAATATAGTCCATCTGCAATCAACAGCACATAGCC 2102
Db 721 TGCTGGGTAGCTTTCACAGACATAAATATAGTCCATCTGCAATCAACAGCACATAGCC 780
Qy 2103 AGGTAACCCACCACTCCCTCGATCTGCAGCCACACAGTGG-GAGTCTGGCTGTAC-C 2160
Db 781 AGGTAACCCACCACTCCCTCGATCTGCAGCCACACAGTGGAGTCTGGCTGTACNC 840
Qy 2161 TTCACAAAGCCACAGAAACCGCCACACA 2187
Db 841 TTCACAAAGCCACAGAAACCGGCCACA 867

RESULT 10
BM044271
LOCUS 603621541F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5446955 5',
DEFINITION mRNA sequence.
ACCESSION BM044271
VERSION BM044271.1 GI:16773538
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC http://imgc.nhl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1930 row: b column: 12
High quality sequence stop: 856.
```

FEATURES

source

Location/Qualifiers

```
1. 877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5446955"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."
```

ORIGIN

```
Query Match 27.5%; Score 811; DB 12; Length 877;
Best Local Similarity 98.9%; Pred. No. 7.4e-93;
Matches 869; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

Qy 1319 GCCAGGCACCAAGCCTAGGGGACCGTACCCGCTCCCTTACACCGACGCGGTTCTGCATG 1378
Db 1 GCCAGGCACCAAGCCTAGGGGACCGTACCCGCTCCCTTACACCGACGCGGTTCTGCATG 60
Qy 1379 AGGCGCAGCGGCTGCTGGCGTGGTCCCATGGGAATACCCCGCACCCCTCATGCGGACCA 1438
Db 61 AGGCGCAGCGGCTGCTGGCGTGGTCCCATGGGAATACCCCGCACCCCTCATGCGGACCA 120
Qy 1439 CCGCTTCCGAGGGGTACACCCCTGCCCGGACGAGGTCCTCCCTCTCTTGGCTCCA 1498
Db 121 CCGCTTCCGAGGGGTACACCCCTGCCCGGACGAGGTCCTCCCTCTCTTGGCTCCA 180
Qy 1499 TCTTCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGG 1558
Db 181 TCTTCATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGG 240
Qy 1559 ATGCAATGAGCGGTTTCAGGAGCATGAGGCTTCTGCTCTCTCTTCTAGGAGCGG 1618
Db 241 ATGCAATGAGCGGTTTCAGGAGCATGAGGCTTCTGCTCTCTCTTCTAGGAGCGG 300
Qy 1619 TCTGCTTGGAGAGGCGCTTGCACCAAGGAGGAGCTCTTCTCTTCTTCAACCATCTAC 1678
Db 301 TCTGCTTGGAGAGGCGCTTGCACCAAGGAGGAGCTCTTCTCTTCTTCAACCATCTAC 360
Qy 1679 AAGCTTCTCTCTGAGAGCGCGTGCCTGCGGACCGTCAAGCTTCAAGCCACCGTCA 1738
Db 361 AAGCTTCTCTCTGAGAGCGCGTGCCTGCGGACCGTCAAGCTTCAAGCCACCGTCA 420
Qy 1739 GTGGCTTTTCAACATTTCCCGAGCTTCCAGCTGCAAGTCCGTCCTGACCTTCACT 1798
Db 421 GTGGCTTTTCAACATTTCCCGAGCTTCCAGCTGCAAGTCCGTCCTGACCTTCACT 480
Qy 1799 CCACCAACGACAGCAGATGAAGAGGCAACTTGAAGTGGTGGTGCAGGACGCGTGC 1858
Db 481 CCACCAACGACAGCAGATGAAGAGGCAACTTGAAGTGGTGGTGCAGGACGCGTGC 540
Qy 1859 CTTCAGCTTCAACAGTGGGATGAGGAGGTTAATGTCTCCAGAGTGTACATGAGGCA 1918
Db 541 CTTCAGCTTCAACAGTGGGATGAGGAGGTTAATGTCTCCAGAGTGTACATGAGGCA 600
Qy 1919 GCCACATTTACACGCTGAGTGTGTTTCCGGAGTGTGTCCCAACGCGCCACAGCTCAC- 1977
Db 601 GCCACATTTACACGCTGAGTGTGTTTCCGGAGTGTGTCCCAACGCGCCACAGCTCAC 660
Qy 1978 TTGACTCATGCTGCTAAGATGACAAACCGCACACCATACACAACTACAGGCGCCACAA 2037
Db 661 TTGACTCATGCTGCTAAGATGACAAACCGCACACCATACACAACTACAGGCGCCACAA 719
Qy 2038 GCAACTGCTGGGTAGCTTTCCACAGACATAATATAGTCCATCTCAATC-ACAAAGCAC 2096
Db 2038 GCAACTGCTGGGTAGCTTTCCACAGACATAATATAGTCCATCTCAATC-ACAAAGCAC
```

Db 720 GCAACTGCTGGTAAAGTTTCCACAGACTTAAATAGTCCATCTGCAATCAACAAGCAC 779

Qy 2097 ATAGCCAGGTAACCAACCACTCCCTGATCTGCAGCCACACGTCGGAGTCTGGCTGT 2156

Db 780 ATAGCCAGGTAACCAACCAATCCCTGGATCTGCAG--CCACAGTGGAGTCTGGCTGT 838

Qy 2157 CACCTTCAAGCCACAG-AAAACGCGCCACACATGTTTCC 2194

Db 839 CACCTTCAAGCCACAGAAAGCGGCACACATGTTCCC 877

RESULT 11

BQ958112 939 bp mRNA linear EST 21-AUG-2002

LOCUS AGENCOURT 10034258 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481143

DEFINITION 5', mRNA sequence.

ACCESSION BQ958112

VERSION BQ958112.1 GI:22373590

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 939)

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTP

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution by: Agencourt Bioscience Corporation

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2662 row: a column: 16

High quality sequence stop: 636.

Location/Qualifiers

1..939

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6481143"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 27.4%; Score 807.8; DB 13; Length 939;

Best Local Similarity 96.8%; Pred. No. 1.8e-92;

Matches 877; Conservative 0; Mismatches 22; Indels 9; Gaps 5;

Qy 645 TGGGGTTTCTTCTCCACGGGGAGCGGTGGAGCGAGCTGAGCAAGTTTACCATGCTGC 704

Db 1 TGGGGTTTCTTCTCC-ACGGGAGCGGTGGAGCGAGCTGAGCAAGTTTACCATGCTGC 59

Qy 705 TCTGCGGACCTGGGCATGGGGAAGCGAAGCGAGGAGCTGATCCAGCGGAGGCCCG 764

Db 60 TCTGCGGACCTGGGCATGGGGAAGCGAAGCGAGGAGCTGATCCAGCGGAGGCCCG 119

Qy 765 GTGCTGTGGTGGACATTCAGGGGACAGAGACGCCCATTCGATCCCTCGCTGTGT 824

Db 120 GTGCTGTGGTGGACATTCAGGGGACAGAGAGCGCCCATTCGATCCCTCGCTGTGT 179

Qy 825 GGCCAGGCGACCTCCAAAGTAGTCTGCTCCCTCTCTTTTGGCCTCCGCTTCTCCTATGA 884

Db 180 GGCCAGGCGACCTCCAAAGTAGTCTGCTCCCTCTCTTTTGGCCTCCGCTTCTCCTATGA 239

Qy 885 GGATAAGGAGTTCAGGCGCGTGGTCCGGGAGCTGGTGGTACCTCTGGGAGTCAGCTC 944

Db 240 GGATAAGGAGTTCAGGCGCGTGGTCCGGGAGCTGGTGGTACCTCTGGGAGTCAGCTC 299

Qy 945 CCAGGGGGTCCAGACTACGAGATGTTCTCTCTGGTTCCTGGGGCCCTGCCAGGCCCCCA 1004

Db 300 CCAGGGGGTCCAGACTACGAGATGTTCTCTCTGGTTCCTGGGGCCCTGCCAGGCCCCCA 359

Qy 1005 CAAGCAGCTCTCCACACAGCTCAGCACTTGGTGGCTTTCACAGTCCGGCAGGTGCAGCA 1064

Db 360 CAAGCAGCTCTCTCCACACAGCTCAGCACTTGGTGGCTTTCACAGTCCGGCAGGTGCAGCA 419

Qy 1065 GCACAGGGGAACCTGGATGCTTCGGGCGCCGACGTCGACCTTCGATGCTTCCTGCT 1124

Db 420 GCACAGGGGAACCTGGATGCTTCGGGCGCCGACGTCGACCTTCGATGCTTCCTGCT 479

Qy 1125 GAAGATGGCAAGGAGGAAACAAACCCAGGACAGAAATTCACCAACAAAGACATGCTGAT 1184

Db 480 GAAGATGGCAAGGAGGAAACAAACCCAGGACAGAAATTCACCAACAAAGACATGCTGAT 539

Qy 1185 GACAGTCATTTATTTGCTGTTTCTGGGACGATGACCGGTGAGCAACCGTTCGCTATAC 1244

Db 540 GACAGTCATTTATTTGCTGTTTCTGGGACGATGACCGGTGAGCAACCGTTCGCTATAC 599

Qy 1245 CCTCTCTCTCTGATGAATACCTCATGTCCTCAAAAGTGGTACGTGAGGAGTGAATCG 1304

Db 600 CCTCTCTCTCTGATGAATACCTCATGTCCTCAAAAGTGGTACGTGAGGAGTGAATCG 659

Qy 1305 GAGAGTGGGGGTGGCCAGGACCAAGCCCTAGGGGACCGTACCCGCCCTCCCTTTACACCGA 1364

Db 660 GAGAGTGGGGGTGGCCAGGACCAAGCCCTAGGGGACCGTACCCGCCCTCCCTTTACACCGA 719

Qy 1365 CGCGGTCTGATGAGCGGAGCGGCTGCTGGCGCTGGTGGCGATGGGAATACCCCGCAC 1424

Db 720 CGCGGNTCTGATGAGCGGAGCGGCTGCTGGCGCTGGTGGCGATGGGAATACCCCGCAC 779

Qy 1425 CCTCATGGGACACCCCGC--TTCGAGGGTACACCTGCCCGCGAGGC----AGGAGTC 1479

Db 780 CCTCATGGGACACCCCGCTTTCCGAGGGTACACCTGCCCGCGAGGCAGGAGTCTT 839

Qy 1480 TTCCCTCTCTTGGCTCCA--TCCTGCATG--ACCCCAACATCTTCAAGCACCCAGAGAG 1536

Db 840 TTCCCTCTCTTGGCTCCA--TCCTGCATG--ACCCCAACATCTTCAAGCACCCAGAGAG 899

Qy 1537 TTCAACCC 1544

Db 900 TTTCACCC 907

RESULT 12

BQ047426

LOCUS 603628712F1 NIH_MGC_40 Homo sapiens linear EST 07-NOV-2001

DEFINITION mRNA sequence.

ACCESSION BQ047426

VERSION BQ047426.1 GI:16776693

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 857)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTP

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1956 row: h column: 04
High quality sequence stop: 817.
Location/Qualifiers
1. .857
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5457075"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 27.2%; Score 801; DB 12; Length 857;
Best Local Similarity 98.5%; Pred. No. 1.4e-91;
Matches 840; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

QY 641 GCCATGGGGTTTCTTCTCCACGGGAGCGGTGGAGGCGCTGAGGAAGTTTACCATGC 700
DB 1 GCCATGGGGTTTCTTCTCCACGGGAGCGGTGGAGGCGCTGAGGAAGTTTACCATGC 60

QY 701 TTGCTCTGGGGACCTGGGCATGGGAAGCGAGAGCGGAGAGCTGATCCAGGGCGAGG 760
DB 61 TTGCTCTGGGGACCTGGGCATGGGAAGCGAGAGCGGAGAGCTGATCCAGGGCGAGG 120

QY 761 CCGGTGTCTGTGGAGACATTCACGGGACAGAGAGCGGCCATTCGATCCCTCCCTGC 820
DB 121 CCGGTGTCTGTGGAGACATTCACGGGACAGAGAGCGGCCATTCGATCCCTCCCTGC 179

QY 821 TGCTGCCAGCGACCTCCACGAGTAGTGTCTCCCTCTCTTTGGCTCCGCTTCTCT 880
DB 180 TGCTGCCAGCGACCTCCACGAGTAGTGTCTCCCTCTCTTTGGCTCCGCTTCTCT 239

QY 881 ATGAGGATAAGGAGTTCCAGGCGGTGTCTCCGGGAGCTGTGTGATACCTGTGGAGTCA 940
DB 240 ATGAGGATAAGGAGTTCCAGGCGGTGTCTCCGGGAGCTGTGTGATACCTGTGGAGTCA 299

QY 941 GCTCCAGGGGGTACAGCTACGAGATGTTCTCTGGTTCTCTGGGCGCCCTGCCAGGCC 1000
DB 300 GCTCCAGGGGGTACAGCTACGAGATGTTCTCTGGTTCTCTGGGCGCCCTGCCAGGCC 359

QY 1001 CCCACAAGCAGCTCTCCACACGCTCAGACACCTTGCTGCTTCCAGTCCGCGAGGTGC 1060
DB 360 CCCACAAGCAGCTCTCCACACGCTCAGACACCTTGCTGCTTCCAGTCCGCGAGGTGC 419

QY 1061 AGCAGCACCAGGGAACTTGGATGCTTCGGGGCCCCGACGTCGTCGATGTCCTTCC 1120
DB 420 AGCAGCACCAGGGAACTTGGATGCTTCGGGGCCCCGACGTCGTCGATGTCCTTCC 479

QY 1121 TGCTGAAGATGGCAGAGGGAGCAAAACCCAGGCACAGATTCACCAACAGAAATGC 1180
DB 480 TGCTGAAGATGGCAGAGGGAGCAAAACCCAGGCACAGATTCACCAACAGAAATGC 539

QY 1181 TGATGACAGTCATTATTGCTGTTGCTGGGACGATGACGGTCAGCACCACCGTCGGCT 1240
DB 540 TGATGACAGTCATTATTGCTGTTGCTGGGACGATGACGGTCAGCACCACCGTCGGCT 599

QY 1241 ATACCTCTCTGCTCCTGATGAATAACCTCATGTCCAAAGTGGGTACGTGAGAGCTGA 1300
DB 600 ATACCTCTCTGCTCCTGATGAATAACCTCATGTCCAAAGTGGGTACGTGAGAGCTGA 659

QY 1301 ATCGGAGCTGGGGCTGGCCAGGCACCAAGCTTAGGGACCGTACCGCCTCCCTTACA 1360
DB 660 ATCGGAGCTGGGGCTGGCCAGGCACCAAGCTTAGGGACCGTACCGCCTCCCTTACA 719

QY 1361 CCGACGGCGTTCTGCATGAGCGGCGAGCGGTGTGTGGCGCTGGTCCCATGGGAATACCC 1420
DB 720 CCGACGGCG-TCTGCATGAGCGGCGAGCGGTGTGTGGCGCTGGTCCCATGGGAATACCC 778

QY 1421 GCACCTCATGGGACCAACCGCTTCCGAGGGTACACCTGCCCGAGGCACGGAGTCT 1480
DB 779 GGAACCTCATGGGACCAACCGCTTCCGAGGGTACA-CCTGCCCGAGGCACGGAGTCT 837

QY 1481 TCCCGCTCCTTGG 1493
DB 838 TCCCGCTCCTTGG 850

RESULT 13
BM044172
LOCUS 6036221408F1 NIH_MGC_40 Homo sapiens cDNA clone linear EST 07-NOV-2001
DEFINITION mRNA sequence.
ACCESSION BM044172
VERSION BM044172.1 GI:16773439
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 804)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999).
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1930 row: o column: 02
High quality sequence stop: 801.
Location/Qualifiers
1. .804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5447257"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN

Query Match 26.6%; Score 783.6; DB 12; Length 804;
Best Local Similarity 99.4%; Pred. No. 2.1e-89;
Matches 797; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1319 GCCAGGACCAAGCTTAGGGACCGTACCGCCTCCCTTACACCGCGGTCTGCATG 1378
DB 1 GCCAGGACCAAGCTTAGGGACCGTACCGCCTCCCTTACACCGCGGTCTGCATG 60

QY 1379 AGCGGACGCGGTGTGTGGCGCTGGTCCCATGGGAATACCCCGCACCTCATCGGACCA 1438

Db 61 AGGCGCAGCGCTGTCGCGCTGTGTCCTATGGAATACCCCGCACCTCATGCGGACCA 120
QY 1439 CCGCTTTCCGAGGATACACCTGTCGCCACGAGGACGAGGCTTCTCCCTCTCTGCTGCTCA 1498
Db 121 CCGCTTTCCGAGGATACACCTGTCGCCACGAGGACGAGGCTTCTCCCTCTCTGCTGCTCA 180
QY 1499 TCTGCTATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTCTCTGG 1558
Db 181 TCTGCTATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTCTCTGG 240
QY 1559 ATGAGATGACCGTTTCCAGAGCATGAGGCTTCTGCGCTTCTCTTAGGGAAGCGTG 1618
Db 241 ATGAGATGACCGTTTCCAGAGCATGAGGCTTCTGCGCTTCTCTTAGGGAAGCGTG 300
QY 1619 TCTGCTTCCGAGGAGGCTGCGCAAGAGGAGCTTCTCTCTTCCACCAACATCTCTAC 1678
Db 301 TCTGCTTCCGAGGAGGCTGCGCAAGAGGAGCTTCTCTCTTCCACCAACATCTCTAC 360
QY 1679 AAGCTTTCTCTGAGAGCGCTGCGCGGACACCTGAGCTTCCGCTTCCACTGACCTTCACT 1738
Db 361 AAGCTTTCTCTGAGAGCGCTGCGCGGACACCTGAGCTTCCGCTTCCACTGACCTTCACT 420
QY 1739 GTGCTTTTCAACATTTCCCGAGCTTCCAGCTTCCAGTCCGCTTCCACTGACCTTCACT 1798
Db 421 GTGCTTTTCAACATTTCCCGAGCTTCCAGCTTCCAGTCCGCTTCCACTGACCTTCACT 480
QY 1799 CCACGAGCAGACCATGAGGAGGAGCACTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTG 1858
Db 481 CCACGAGCAGACCATGAGGAGGAGCACTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTG 540
QY 1859 CTCAGGCTCAACAGTGGGATGACAGGCTTAAATGCTCCAGAGTGTACACTGCGAGCA 1918
Db 541 CTCAGGCTCAACAGTGGGATGACAGGCTTAAATGCTCCAGAGTGTACACTGCGAGCA 600
QY 1919 GCACATTTACAGCTGAGTGTGTTTCCGAGTGTGCTCCAGGCGCCACAGCTCACT 1978
Db 601 GCACATTTACAGCTGAGTGTGTTTCCGAGTGTGCTCCAGGCGCCACAGCTCACT 660
QY 1979 TGACTATGCTGATAGTGCACAGGACACCCATACAACTCAAGGGCCACAAAG 2038
Db 661 TGACTATGCTGATAGTGCACAGGACACCCATACAACTCAAGGGCCACAAAG 720
QY 2039 CAAGTCTGGTGTAGTGTGCTTCCAGACATAAATATAGTCCATCTGCAATCACAGCACAT 2098
Db 721 CAAGTCTGGTGTAGTGTGCTTCCAGACAT-ACTATAGTCCATCTGCAATCACAGCACAT 779
QY 2099 AGCAGGTAAACCACTCC 2120
Db 780 AGCAGGTAAACCACTCC 801

RESULT 14
LOCUS BU543186
DEFINITION AGENCOURT 10338701 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6575194
5', mRNA sequence.
ACCESSION BU543186
VERSION BU543186.1 GI:22853669
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2771 row: 1 column: 10
High quality sequence stop: 703.
Location/Qualifiers

FEATURES
source

1..877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6575194"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 26.0%; Score 764.8; DB 13; Length 877;
Best Local Similarity 97.8%; Pred. No. 4.8e-87;
Matches 808; Conservative 0; Mismatches 12; Indels 6; Gaps 3;
QY 645 TGGGGTTTCTTCTTCTCAACGGGGAGCGGTGGAGGAGCTGAGGAAGTTTACCATGCTTGC 704
Db 1 TGGGGTTTCTTCTTCTCC-ACGGGGAGCGGTGGAGGAGCTGAGGAAGTTTACCATGCTTGC 59
QY 705 TCTCGGGAGCTTGGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 764
Db 60 TCTCGGGAGCTTGGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 119
QY 765 GTGTCTGTGGAGACATTCAGGGGACAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 824
Db 120 GTGTCTGTGGAGACATTCAGGGGACAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 179
QY 825 GGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 884
Db 180 GGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 239
QY 885 GGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 944
Db 240 GGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 299
QY 945 CCAGGGGGTCCAGACCTACGAGATGTTCTCTGGTTTCTCGGGCCCTTGCAGGCGCCCA 1004
Db 300 CCAGGGGGTCCAGACCTACGAGATGTTCTCTGGTTTCTCGGGCCCTTGCAGGCGCCCA 359
QY 1005 CAAGCAGCTCTCCACGAGCTGAGCAGCTTGGCTGCTTCCAGTCCGAGGAGGAGGAGGAG 1064
Db 360 CAAGCAGCTCTCTCCACGAGCTGAGCAGCTTGGCTGCTTCCAGTCCGAGGAGGAGGAGG 419
QY 1065 GCACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1124
Db 420 GCACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 479
QY 1125 GAAGATGCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1184
Db 480 GAAGATGCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539
QY 1185 GACAGTCAATTTATTTGCTGTTTGGGAGCATGAGTCCAGCAGCAGGTCGGGTATAC 1244
Db 540 GACAGTCAATTTATTTGCTGTTTGGGAGCATGAGTCCAGCAGCAGGTCGGGTATAC 599
QY 1245 CTTCTGCTCTCTGATGAATACCTTCATGTCCTCAAAAGGGGTACCTGAGGAGCTGAATCG 1304
Db 600 CTTCTGCTCTCTGATGAATACCTTCATGTCCTCAAAAGGGGTACCTGAGGAGCTGAATCG 659

1305 GGAGCTGGGGCTGGCGAGGACCAAGCTAGGAGCGGTACCGCTCCCTTACACGA 1364
 660 GGAGCTGGGGCTGGCGAGGACCAAGCTAGGAGCGGTACCGCTCCCTTACACGA 719
 1365 GGCGCTTCTGATGAGGCGAGCGGCTGCTGGCGCTGG-TGCCATGGGAATACCCGCA 1423
 720 GGCGCTTCTGATGAGGCGAGCGGCTGCTGGCGCTGGTGGCCATGGGAATACCCGCA 779
 1424 CCCTCATGCGG-ACCACCGCTTCCGAGGTACACCTGGCCC 1465
 780 CCCTCTGGGGACCAACCGGCTTCCGAGGGGTACACCTGGCCC 825

RESULT 15
 LOCUS BX422464
 DEFINITION BX422464 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
 CS0DM006YC04 5-PRIME, mRNA sequence.
 ACCESSION BX422464
 VERSION BX422464.1 GI:30769754
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 969)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9634.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DM006BB02QP1&cluster=9634.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DM006BB02QP1.
 Location/Qualifiers
 1..969

FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DM006YC04"
 /tissue_type="FETAL LIVER"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL LIVER"
 /note="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

ORIGIN
 Query Match 24.3%; Score 716.2; DB 13; Length 969;
 Best Local Similarity 89.4%; Pred. No. 6.1e-81;
 Matches 779; Conservative 30; Mismatches 53; Indels 9; Gaps 4;
 282 GAGAGGAGCGGACCTGGAGATGGAGGCGCGGACCTGGCGCTGCTGCTGGCGCT 341
 76 GAGAGGAGCGGACCTGGAGATGGAGGCGCGGACCTGGCGCTGCTGCTGGCGCT 135
 342 GGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 401
 136 GGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 195
 402 CCGCGGCGGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
 196 CCGCGGCGGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 255

Qy 462 GTATTGAGGCTCATGGGCTGAGTAAGAGTACGACCGGTGTTCACCATCTACCTGGG 521
 Db 256 GTATTGAGGCTCATGGGCTGAGTAAGAGTACGACCGGTGTTCACCATCTACCTGGG 315
 Qy 522 ACCGTGCGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 581
 Db 316 ACCGTGCGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 375
 Qy 582 TCAGGCTGAGGAGTTCAGCGGCGGGAACCGTACGATGCTGGAAGGAGCTTTTGATGG 641
 Db 376 TCAGGCTGAGGAGTTCAGCGGCGGGAACCGTACGATGCTGGAAGGAGCTTTTGATGG 435
 Qy 642 CCATGGGCTTTTCTTCCAAACGGGAGCGGTGGAGGAGCTGAGGAAGTTTACCATGCT 701
 Db 436 CCATGGGCTTTTCTTCCAAACGGGAGCGGTGGAGGAGCTGAGGAAGTTTACCATGCT 495
 Qy 702 TGCTCTGGGACCTGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGG 761
 Db 496 TGCTCTGGGACCTGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGG 555
 Qy 762 CCGGTGCTGCTGGAGACATTCAGGGGACAGAGAGCGCCATTCGATCCCTCCCTGCT 821
 Db 556 CCGGTGCTGCTGGAGACATTCAGGGGACAGAGAGCGCCATTCGATCCCTCCCTGCT 615
 Qy 822 GCTGCCCGAGGCTCCTCCAAACGTAGTCTGCTCCCTCTCTTTGGCCTCCGCTTCTCCTA 881
 Db 616 GCTGCCCGAGGCTCCTCCAAACGTAGTCTGCTCCCTCTCTTTGGCCTCCGCTTCTCCTA 675
 Qy 882 TGAGGATAGGAGTTCCAGGCGGTGGTCCGGGAGCTGGTGTACCTGCTGGGAGTCAG 941
 Db 676 TGAGGATAGGAGTTCCAGGCGGTGGTCCGGGAGCTGGTGTACCTGCTGGGAGTCAG 735
 Qy 942 CTCCCGAGGCTCAGACCTACGAGATGTTCTCTGTTTCTTGGGCGGCTCCAGGGCCC 1001
 Db 736 CTCCCGAGGCTCAGACCTACGAGATGTTCTCTGTTTCTTGGGCGGCTCCAGGGCCC 795
 Qy 1002 CCACAAAGCAGCTCTCCACCGCTCAGACCTTGGCTGCTTTCACAGTCCGGCAGGTGA 1061
 Db 796 C-AAAAGACTCTCCACCAAGKCA-SMCTTKGGCGSKKCAAAARKGSGAAGKMRCA 852
 Qy 1062 GCAGCACCAGGGGAACTGGATGCTTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCTTCT 1121
 Db 853 GCA----CAAGGGRAACTGGATGCTT--GGGCGGCGGCGGCGGCGGCGGCGGCGG 906
 1122 GCTGAAGATGGCAGAGGAGGAACAAAACCCA 1152
 907 CGAAAAGGCAAAAGAAAGAAAACAAAGCMA 937

Search completed: September 17, 2004, 18:58:19
 Job time : 7181.05 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 17, 2004, 19:05:01 ; Search time 5865 Seconds
(without alignments)
3724.619 Million cell updates/sec

Title: US-10-669-693-2
Perfect score: 2615
Sequence: 1 MEAGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10669693/runat_15092004_101310_5847/app_query.fasta_1.647
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10669693 @CGN 1 1 3731 @runat_15092004_101310_5847 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

GenEmbl:
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.in.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rat.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2615	100.0	1515	6	AX591678 Sequence
2	2615	100.0	2261	6	AX780433 Sequence
3	2615	100.0	2261	9	AF335278 Homo sapi
4	2615	100.0	2604	6	AX430320 Sequence
5	2615	100.0	2619	9	BC033691 Homo sapi
6	2615	100.0	2620	6	AX327331 Sequence
7	2615	100.0	2623	6	AX463930 Sequence
8	2615	100.0	2626	9	AY358603 Homo sapi
9	2610	99.8	2604	6	AX552230 Sequence
10	2527	96.6	2513	6	AX876954 Sequence
11	2527	96.6	2513	6	BD156406 Primer fo
12	2527	96.6	2513	9	AK027605 Homo sapi
13	1628.5	62.3	2114	6	AX552231 Sequence
14	1313	50.2	2026	4	RABF450BX
15	1313	50.2	2028	4	S64259
16	1301	49.8	2045	4	RABCYCP2B
17	1301	49.8	2743	4	AB052256 Sus scro
18	1300	49.7	2065	4	RABCYCP2C
19	1300	49.7	2081	10	BC034202 Mus muscu
20	1299.5	49.7	1795	10	AF128849 Mus muscu
21	1299.5	49.7	1901	10	BC060973 Mus muscu
22	1299.5	49.7	2079	4	RABCYCP2A
23	1289	49.3	1802	4	RABP450CA
24	1289	49.3	1567	10	RATCP450
25	1289	49.3	1567	10	RATCP45Z
26	1286.5	49.2	2004	6	AX477702
27	1280.5	49.0	1822	10	MUSTH16AB
28	1262	48.3	2642	4	DOGPA591IB
29	1261.5	48.2	1717	10	D86952
30	1257	48.1	1480	6	E63778
31	1255	48.0	1988	10	AY250712 Mesocricet
32	1245.5	47.6	1853	10	L81171 Mus musculu
33	1245.5	47.6	1886	10	BC058222 Mus muscu
34	1244	47.6	1853	10	MUSTH16A
35	1243.5	47.6	1857	10	RATCVT45XX
36	1242.5	47.5	1781	4	AB052255
37	1239	47.4	3045	6	AX018551
38	1239	47.4	3045	6	AX018627
39	1239	47.4	3045	6	BD136285
40	1238	47.3	1476	6	E10616
41	1238	47.3	1476	6	E10861
42	1238	47.3	1476	6	AR399329
43	1238	47.3	3045	6	AX332736
44	1238	47.3	3045	6	AX409672
45	1238	47.3	3045	9	HUMCYP2BB

ALIGNMENTS

RESULT 1

Mon Sep 20 09:09:25 2004

AX591678 1515 bp DNA linear PAT 27-JAN-2003

LOCUS AX591678 Sequence 39 from Patent WO0246409.

DEFINITION AX591678

ACCESSION AX591678.1 GI:27950054

VERSION

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS Guo, X., Li, L., Patturajan, M., Shimkets, R. A., Casman, S. J., Malyanar, U. M., Tchernev, V. T., Vernet, C. A., Spytek, K. A., Shenoy, S. G., Alsobrook, J. P., Edinger, S., Peyman, J. A., Stone, D. J., Ellerman, K., Gangolli, E. A., Boldog, F. L., Colman, S. D., Eisen, A. J., Liu, X., Padigaru, M., Spaderna, S. K. and Zerhusen, B. D.

TITLE Proteins and nucleic acids encoding same

JOURNAL Patent: WO 0246409-A 39 13-JUN-2002;

Curagen Corporation (US)

FEATURES

Location/Qualifiers

1..1515

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: Length: 1515

Score: 2615.00 Matches: 504

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: Gaps: 0

US-10-669-693-2 (1-504) x AX591678 (1-1515)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20

Db 1 ATGGAGGCGACCGGACCTGGGGCTGCTGGCGCTGGCGCTGGCGCTGGCGCTG 60

QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40

Db 61 CTGGCGCTGTCGGGACCGAGGCGCCGAGGCGACCTGCCCGCGGGCGCCGCGCTACCA 120

QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyValAlaLeuTyrSerGlyLeuMetArgLeu 60

Db 121 CTGCTGGGAAACCTCTGAGCTACGGCCCGGGCGCTGTATTACGGGCTCATGGCGCTG 180

QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValVal 80

Db 181 AGTAAGAAGTACGACCGGTTTACCATCTACCTGGGACCGGCGCTGTGGTGGTC 240

QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100

Db 241 CTGGTTGGGCGAGAGGCTGTGGGAGGCGCTGGAGGTACGCTGAGGAGTTACGGCGC 300

QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120

Db 301 CGGGGAACCTAGCGATGCTGGAAGGACATTTGATGGCCATGGGGTTTCTTCTCCAAC 360

QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140

Db 361 GGGGAGCGGTGGAGGAGCTGAGGAAGTTTACCATGCTTGTCTGCGGGACCTGGGCATG 420

QY 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160

Db 421 GGGGAAGCGAGAGCGAGGAGCTGATCCAGCGGAGGCGCGGCTGTCTGTGGAGACATTC 480

QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180

Db 481 CAGGGGACAGAGGAGCGCCATTCGATCCCTCCCTGCTGCTGGCGCCAGGCCACCTCAAC 540

QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200

Db 541 GTAGTCTGCTCCCTCTCTTTGGCCTCGCTTCTCTATGAGGATAAGGAGTTCAGGCC 600

QY 201 ValValArgAlaAlaGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220

Db 601 GTGGTCCGGGCGAGCTGGTGTGTACCTGCTGGAGTCAGCTCCCGGGGGGTGACAGCTTAC 660

QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240

Db 661 GAGATGTTCTCTCTGGTTCCTGGGCCCTCTGCAGGCCCCCAGAGCAGCTCTCTCCACCAC 720

QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260

Db 721 GTGAGCAGCTTGGTGGCTTCCAGCTCGGCGAGGTGCAGCAGCACACGAGGAACTGGAT 780

QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280

Db 781 GCTTCGGGCGCCGACGCTGACCTTGTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 840

QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300

Db 841 CAAACCCAGGCGACAGAAATTCACCAACAGAAATGCTGATGACAGTCAATTTATTGCTG 900

QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320

Db 901 TTTGCTGGGACGATGACGCTGACGACCGCTGGCTATACCTCTCTCTGCTGATGAAA 960

QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340

Db 961 TACCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAG 1020

QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360

Db 1021 GCACCAAGCCTAGGGGACGCTACCGCTCCCTTACACGACGCGCTTCTGCTGATGGCG 1080

QY 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrArg 380

Db 1081 CAGCGGCTGCTGGCGTGGTGGCCATGGGAATACCCCGCACCTCATGCGGACCCCGC 1140

QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400

Db 1141 TTTCCGAGGGTACACCTGCTGCCCGAGGCGACGGAGGTCTTCCCTCTCTGGCTCCATCT 1200

QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420

Db 1201 CATGACCCCAACATCTTCAGGACCCAGAGAGTTCAACCCAGACCGTTTCTGTGATGCA 1260

QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440

Db 1261 GATGAGCGGTTTCAGGAAGCATGAGGCGTTCTCTGCCCTCTCTCTTAGGAGAGCGTCT 1320

QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460

Db 1321 CTTTGGAGAGGCGCTGGCAAAAGCGAGCTCTTCTCTTCTTACCCACCTCTTACAGCC 1380

QY 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480

Db 1381 TTTCTCTCGAGAGCGCGTGGCGCGCGGACACCTGAGCCTCAAGCCACCGTCAGTGGC 1440

QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500

Db 1441 CTTTTCACCATTTCCCGCAGCTTCCAGCTCGCAAGTCGCTCCCTGCTGCTGCTGCTGCT 1500

QY 501 ThrGlnThrArg 504

Db 1501 ACGGAGACCCAGA 1512

RESULT 2

AX780433 2261 bp DNA linear PAT 14-JUL-2003

LOCUS AX780433

DEFINITION Sequence 2590 from Patent WO03039443.

ACCESSION AX780433

VERSION AX780433.1 GI:32697427

[illegible]

DEFINITION Homo sapiens cytochrome P450, family 2, subfamily S, polypeptide 1, mRNA (cdna clone MGC:44853 IMAGE:5212609), complete cds.

ACCESSION BC033691
VERSION BC033691.1 GI:21707057
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2619)
AUTHORS Strausberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Maras, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheet, R.E., Brownstein, M.J., Uesdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257
PUBMED 12477932

REFERENCE 2 (bases 1 to 2619)
AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: mgc@nci.nih.gov
Email: cgabs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov

AKhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maki, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W., Teurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 68 Row: 0 Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 20522238.

FEATURES
Location/Qualifiers
1..2619

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:44853 IMAGE:5212609"

/tissue type="Blood, adult leukocytes"

/clone_lib="NIH_MGC_118"

gene

/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2619
/gene="CYP2S1"
/db_xref="LocusID:29785"
18..1532
/codon_start=1
/product="cytochrome P450, family 2, subfamily S, polypeptide 1"
/protein_id="AAH33691.1"
/db_xref="GI:21707058"
/db_xref="LocusID:29785"
/translation="MEATGTWALLLALLLLLTALSGTRARGHLPGPTPLPGLN LLQRLPGLYSGLMRLSKYGVFTYILGPRVVLVVGQAEVREALGGQAEFGRRG TLMLEGTFDGTVFSGNGRWOLRKFETMLADLGMKREGEELIQAEARCLVETI QCTGRRPDPSSLAAQKATSNVCSLLFLGRFSYEDKEFOAVRAAGTLLGVSSQGGQ TYEWSFLRPLPDPGKOLLHHVSTLAFTVQVQHQGNLDASGPARDLVDAFLKRM AQEQNPCTEFTNKNMLVIYLLFAGTMTVSTVGYTLLLLMKYPHVQKVRBELNR ELGAGQPSLGDRTLRPYTDAVLEAQLALVPMGIPKTLMLRTRFRGYTLPGQTEV FPLIGSILHDPNIFKHPFENPDRFLADGRFKEAFPLFSGLKRVCLGEGLAKAEL FLFFTILQAFSLSPCPDPTLSLKLPTVSGLFNIPPAFQLOVRPTDLHSTTQR"

CDS

misc_feature

177..1493
/note="p450; Region: Cytochrome P450. Cytochrome P450s are involved in the oxidative degradation of various compounds. Particularly well known for their role in the degradation of environmental toxins and mutagens. Structure is mostly alpha, and binds a heme cofactor"
/db_xref="CDD:pfam00067"

ORIGIN

Alignment Scores:

Pred. No.: 5.3e-201 Length: 2619
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-669-693-2 (1-504) x BC033691 (1-2619)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuLeuLeuLeuLeuLeuLeuThr 20
DB 18 ATGGAGCGGACCGGCACCTGGCGCTGCTGGCGCTGGCGCTGGCGCTGGCGCTG 77
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40
DB 78 CTGGCGCTGTCGGGACCGAGGCGCCGAGGCCACCTGCCCGCGGCCCGCCGCTACCA 137
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
DB 138 CTGCTGGAAACCTCTCTCAGCTACGCGCGCGCGCGCTGTATTTCAGGGCTCATGCGGCTG 197
QY 61 SerIysIysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValValVal 80
DB 198 AGTAAGAAGTACCGAGCGGCTTACCATCTACTGGGACCCCTGGCGCGCTGTGGTGTG 257
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
DB 258 CTGTTGGGAGGAGGCTGTGCGGAGGCCCTGGAGGCTGAGGCTGAGGAGTTTCCAGCGGC 317
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
DB 318 CGGGAAACCTAGCATGCTGGAAGGACTTTTGTATGCCCATGGGGTTTCTTCTCCAC 377
QY 121 GlyGluArgTrpArgGlnLeuArgGlyPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
DB 378 GGGAGCGGTGGAGGAGGCTGAGGAAGTTTACCATTGCTGCTGCGGGACCTGGCGCATG 437
QY 141 GlyIysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
DB 438 GGGAGCGAGAGCGAGGAGCTGATCCAGCGGAGGCCCGCGTGTCTGTGGTGGAGACATTC 497
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180

581 GTAGTGTGCTCCCTCTCTTTGGGCTCGCTTCTCTATGAGGATAAGGAGTTCAGGCC 640
 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
 641 GTGTCGGGAGCTGGTGGTACCTCTGTGGAGTCACTCCAGGGGGTTCAGACCTAC 700
 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
 701 GAGATGTTCTCTGCTGCTGCTGGGCTCGCAGGCCGCCCAAGCAGCTCTCTCCACCAC 760
 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnHisGlnGlnValAsnLeuAsp 260
 761 GTCAGACCTTGGTGGCTTTCACATGTCGGCAGGTGCAGACGACACAGGGGACCTGGAT 820
 261 AlaSerGlyProAlaArgAspLeuValAlaPheLeuLeuLysMetAlaGlnGlu 280
 821 GCTTCGGGCCCGCCAGCTGACCTTGTGATGCTTCTCTGCTGAAGATGACAGGAGAA 880
 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
 881 CAAAACCCAGGCACAGAAATTCACCAACAGAACATGCTGATGACAGTCAATTTATTTGCTG 940
 301 PheAlaGlyThrMetThrValSerThrValGlyTyrThrLeuLeuLeuMetLys 320
 941 TTTGCTGGGACGATGACGTCAGCACCCAGCTCGGCTATACCTCTCTCTCTGATGAAA 1000
 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGln 340
 1001 TACCCTCATGTCCAAAGTGGGTACGTGAGGAGTGAATCGGGAGCTGGGGGTGCCAG 1060
 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
 1061 GCACCAAGCTTAGGGACCGTACCCGCTTCCCTTACACGACGCGGTTCTGATGAGGG 1120
 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
 1121 CAGCGGCTGTGGCGTGTGGTCCCATGGGAAATACCCGACCCCTCATGGGACCAACCCG 1180
 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
 1181 TTCGAGGTATACCTGTCGCCAGGACGAGGAGTCTTCCCTCTCTTGGCTCATCTCTG 1240
 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
 1241 CATGACCCCAACATCTTCAAGCACCCAGAGAGTCAACCCAGACCGTTTCTCTGGATGCA 1300
 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
 1301 GATGGACGTTTCAGGAAGCATGAGGCTTCTGCTCTCTCTTCCCTTAGGAGCGTGTCTG 1360
 441 LeuGlyGlyGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
 1361 CTTGGAGAGGCTGGCAAGCGGAGTCTTCTCTCTTCTTCTTCCACCATCTCTACAGCC 1420
 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
 1421 TTTCTCCCTGAGAGCCGTCGCGCGGACACCCCTGAGCCTCAAGCCACCGTCAGTGGC 1480
 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
 1481 CTTTTCACATTTCCCCAGCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1540
 501 ThrGlnThrArg 504
 1541 ACGCAGACCAGA 1552

RESULT 7
 AX463930
 LOCUS
 DEFINITION Sequence 63 from Patent WO0140466.
 AX463930
 ACCESSION
 VERSION AX463930.1 GI:21898982
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS
 Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,
 Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
 Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
 Wood,W.L. and Zhang,Z.,
 TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
 same
 JOURNAL Patent: WO 0140466-A 63 07-JUN-2001;
 Genentech Inc. (US)
 FEATURES
 source
 1. 2623
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,31e-201 Length: 2623
 Score: 2e15.00 Matches: 504
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-10-669-693-2 (1-504) x AX463930 (1-2623)
 QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuLeuLeuThr 20
 Db 41 ATGGAGGCGACCGCACCTGGCGCTGCTGCTGGCGCTGGCGCTGCTGCTGCTGCTGCG 100
 QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProThrProLeuPro 40
 Db 101 CTGGCGCTGTCCGGGACCGAGGCGCCGAGGCGCCCTGCCCCCGGGCCGACGCGCTACCA 160
 QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
 Db 161 CTGCTGGGAAACCTCTCGAGCTACGCGCGGGGCGCTGTATTTCAGGGCTCATGCGGCTG 220
 QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValValVal 80
 Db 221 AGTAAGATGACGACCGCGGTTCACCACTACTCGGACCCCTGGCGGCGCTGTGGTGGTC 280
 QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
 Db 281 CTGGTTGGGCGAGGAGCTGTGGGGAGGCGCTGGGAGGCTGAGGCTGAGGAGTTTTCAGCGGC 340
 QY 101 ArgGlyThrValAlaMetLeuGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
 Db 341 CGGGGAACCGTACCGATGCTGGAGGAGACTTTTGTATGGCCATGGGTTTTTCTTCTTCCAAC 400
 QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
 Db 401 GGGGAGCGGTGGAGGAGCTGAGGAAGTTTACCATGCTTCTGCGGGACCTGGGCGATG 460
 QY 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
 Db 461 GGGGAACGGAAGCGGAGGAGCTGATCCAGCGGAGGCGCGGCTGTCTGTGGGAGACATTC 520
 QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180
 Db 521 CAGGGACAGAGGACCGCCATTCGATCCCTCTCTGCTGCTGCGCCAGGCCACCTCCAAC 580
 QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
 Db 581 GTAGTGTCTCTCTCTTGGCGCTTCTCTATGAGGATAAGGAGTTCAGGCC 640
 QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
 Db 641 GTGGTCCGGGACGCTGTGTGTGTACCTCTGCTGGAGTCACTCCAGGGGGGTTCAGACCTAC 700

Qy	221	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis	240
Db	701	GAGATGTTCTCCTGGTTCCTGGCCCTCGCAGGCCCCCAACAGAGCTCCTCCACCAC	760
Qy	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
Db	761	GTCAGCACCTTGGTGCTTCACAGTCCGGCAGGTGCAGCAGCACAGGGGAACCTGGAT	820
Qy	261	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu	280
Db	821	GCITCGGGCCCCGACGCTTGTGTGATGCCCTTCCTGCTGAAGATGGCAGAGAGAA	880
Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
Db	881	CAAAACCCAGGCACAGAAATTCACCAACAGAACATGCTGATGACAGTCAATTTTTCGCTG	940
Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuLeuMetLys	320
Db	941	TTTGTCTGGGACGATACCGTCAGCACACCGTTCGGCTATACCTTCCTGCTCTGTATGAAA	1000
Qy	321	TyrProHisValGlnLysTyrValArgGluLeuLeuAsnArgGluLeuGlyAlaGlyGln	340
Db	1001	TACCCATATGTCAAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAG	1060
Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1061	GCACCAAGCCTAGGGGACCGTACCCGCTCCCTTACACCGACGCGTTCCTGCAATGAGCG	1120
Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
Db	1121	CAGCGCTGTGGCGTGTGGTCCCATGGGAATACCCCGCACCTCATGGGACACCCCGC	1180
Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1181	TTCCGAGGCTACACCTGCCCGCCAGGGCAGCGAGTCTTCGCCCTCTTGCCTCCATCTG	1240
Qy	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1241	CATGACCCCAACATCTTCAAGCACCCAGAAAGATTCAACCCAGACCGTTTCTCTGGATGCA	1300
Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1301	GATGACCGTTCAGGAAGCATGAGGCTTCCTGCCCTTCTCTTGGGAAGCGTGCTCG	1360
Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla	460
Db	1361	CTTGAGAGGGCTGGCAAAAGCGGAGCTCTTCTCTTCTTCCACCACCTCTACAGCC	1420
Qy	461	PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1421	TTCTCCCTGGAGGCCGCGTGGCCGGCGGACACCTGAGCTTCAAGCCACCGCTCAGTGGC	1480
Qy	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValAtgProThrAspLeuHisSerThr	500
Db	1481	CTTTTCAACATTCCTCCAGCCTTCACAGTGCAGTCCGATCCCTCCACCTGACCTTCACTCCACC	1540
Qy	501	ThrGlnThrArg	504
Db	1541	ACGCAGACCCAGA	1552

RESULT 8	
AY358603	
LOCUS	
DEFINITION	AY358603 Homo sapiens clone DNA76395 Cytochrome p450 (UNQ891) mRNA, complete cds.
ACCESSION	AY358603
VERSION	AY358603.1 GI:37182327
KEYWORDS	FLI_CDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2626)
Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B.,
Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,
Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,
Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.,
Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V.,
Stinson,J., Vagts,A., Vanden,R., Watanabe,C., Wiedand,D., Woods,K.,
Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,
Goddard,A., Wood,W.I. and Godowski,P.
The Secreted Protein Discovery Initiative (SPDI), a Large-Scale
Effort to Identify Novel Human Secreted and Transmembrane Proteins:
A Bioinformatics Assessment
Genome Res. 13 (10), 2265-2270 (2003)
2 (bases 1 to 2626)
Clark,H.F.
Direct Submission
Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,
Inc., 1 DNA Way, South San Francisco, CA 94080, USA.
Location/Qualifiers
1. .2626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DNA76395"
1. .2626
/locus_tag="UNQ891"
44. 1558
/locus_tag="UNQ891"
/note="PRO1906"
/codon_start=1
/product="Cytochrome p450"
/protein_id="AAQ88966.1"
/db_xref="GI:37182328"
/translation="MEATGTWALLALALLLTALSGTRARGLPGPTPLFLGN
LQLRPGALYGLMLRSKYGVPVTVLGPWRPVVLVGQAVREALGQAEFGSRG
TVAMLEGDFDGVFFSNGERNQLRKFTLMALRDNLGMKRGCEELIQAEARCLVTF
QCTEGRFPDPSLLLAQATSNVVCSLFLGLRFSYEDKEFOAVRAGGTLGVSSQGG
TVEMFSWFLRPLPGPKQLLHVSTLAATFVQVQHQNLDSAGDAGTLVDAPFLKM
AEQEQNPCTFTNKNMLTVIYLLFAGTMTVSTVGYTLLLLMKYPHVQKVRBELRM
ELGACQASLGDRTRLPYDAVLEHAQRLALVPMGIPRTLMRTTFRGYTLPGQGTVE
FLLGLSI LHDNPNI FKHPPEYDFDLADGRFRKHEAFPLFSLGKRVCLGEGLAKEA
FLFTFTILLQAFLSESPCPDITLSLXPTVSGLFNI PPAFOI OVRPTD LHSYTPP"

ORIGIN

```

Alignment Scores:
Pred. No.:          5.32e-201      Length:      2626
Score:              2615.00      Matches:     504
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches:   0
Query Match:        100.00%      Indels:       0
DB:                  9            Gaps:         0

US-10-669-693-2 (1-504) x AY358603 (1-2626)

```

[illegible]

Db	1364	CTTGGAGAGGCGCTGGCAAAAGCGAGAGCTCTTCTCTCTTCTTCCACCACCATCTACAGCC	1423
Qy	461	PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1424	TTCTCCCTGGAGAGCGCTGCGCGCGGACACCTGAGCCTCAAGCCACCGTCAGTGGC	1483
Qy	481	LeuPheAsnIleProProAlaPheGlnLeuInValArgProThrAspLeuHisSerThr	500
Db	1484	CTTTTCAACATTCGCCAGCCTTCAGCTGCAAGTCCGCTCCACTGACCTTCACCTCCACC	1543
Qy	501	ThrGlnThrArg	504
Db	1544	ACGAGACCCAGA	1555
RESULT 9			
AX552230		2604 bp	DNA linear PAT 27-NOV-2002
LOCUS	AX552230	Sequence 139 from Patent WO0162927.	
DEFINITION	AX552230		
ACCESSION	AX552230.1	GI:25896468	
VERSION			
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 Banville, S.C., Greenawalt, L.B., Lincoln, S.E., Stockdreher, T.K., Ameshey, S., Chang, S.C., Chen, W., D'Sa, S.A., Dam, T.C., Liu, T.F., Rosen, B.H., Russo, F.D., Spiro, P.A., Bradley, D.L., Chen, A., Cohen, H.J., Daffo, A., Daniels, S.E., Dufour, G.E., Flores, V., Fong, W.T., Hodgson, D.M., Jackson, S., Jones, A.L., Panzer, S., Roseberry, A.M., Shah, P., Wright, R.J., Yap, P.E., Yu, J.Y., Bratcher, S.R., Chalup, M.S., Dahl, C.R. and Hillman, J.L.	
TITLE		Polypeptides and corresponding polynucleotides for diagnostics and therapeutics	
JOURNAL		Patent: WO 0162927-A 139 30-AUG-2001;	
FEATURES		Incyte Genomics, Inc. (US)	
SOURCE		Location/Qualifiers	
		1..2604	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
		/note="Incyte ID No: LG:337818.2:2000FEB18"	
ORIGIN			
Alignment Scores:		Length: 2604	
Pred. No.:	1..33e-200	Matches: 503	
Score:	2610.00	Conservative: 0	
Percent Similarity:	99.80%	Mismatches: 1	
Best Local Similarity:	99.80%	Indels: 0	
Query Match:	99.81%	Gaps: 0	
DB:			
US-10-669-693-2 (1-504) x AX552230 (1-2604)			
Qy	1	MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr	20
Db	40	ATGAGGCGGACCGGACCTGGCGCTGCTGCTGGCGTGGCGTCTCTGCTGCTGCTGCTG	99
Qy	21	LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro	40
Db	100	CTGGCGCTGCTGGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	159
Qy	41	LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu	60
Db	160	CTGCTGGGAAACCTCTGAGCTAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	219
Qy	61	SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValVal	80
Db	220	AGTAAGAAGTACGAGCGCGGTTCACCATCTACCTGGGACCGCTGGCGGCGGCGGCG	279
Qy	81	LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly	100

```
Db 280 CTGGTGGGACAGAGCTGTGCGGAGGCGCCCTGGGAGGTTCAGGAGTTTCAGCGGC 339
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 340 CGGGAACCGTAGCATGCTGGAAGGACTTTTGATGGCCATGGGGTTTCTTCTCCAAC 399
QY 121 GlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db 400 GGGGAGCGGTGAGGACAGCTGAGGAAGTTTACCATGCTTGTCTCTCGGGAGCTGGGCATG 459
QY 141 GlyLysArgGluGlyGluGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db 460 GGGGAAGCAGAGGCGAGGAGCTGATCCAGGCGGAGGCGCGGTGTCTGCTGGAGACATTC 519
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180
Db 520 CAGGGACAGAGAGCGCCCATTCATCCCTCTGCTGCTGGGCCAGGCCACCTCCAAC 579
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db 580 GTAGTCTGCTCCCTCTTGGGCTCGGCTTCTCCATGAGGATAGGAGTTCCAGGCC 639
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr 220
Db 640 GTGTCGCGGAGCTGTGTGTACCTCTGCTGGAGTTCAGTCCAGGGGGGTTCAGACCTAC 699
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
Db 700 GAGATGTTCTCTGTTCTGCGGCCCTGCGAGGCCCCACAGCAGCTCTCCACAC 759
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db 760 GTCAGCACCTTGGCTGCTTACAGTCCGCGAGTGCAGCAGCACACAGGGAACCTGGAT 819
QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
Db 820 GCTTCGGGCCCCGAGCTGTGTCGATGCTTCTGCTGAGATGGCACAGGAGAA 879
QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db 880 CAAACCCAGGCACAGATTCACCAACAGAACATGCTGATGACAGTCATTTATTGTGCTG 939
QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
Db 940 TTTGCTGGGACGATGCGGTGAGCCACCGCTCGGCTATACCTCTCTCTCTGATGAAA 999
QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
Db 1000 TACCCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGTGGCCAG 1059
QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db 1060 GCACCAAGCCTAGGGACCGTACCCGCTCCCTTACCGACGCGGTTCGTGATGAGCGC 1119
QY 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
Db 1120 CAGCGCTGCTGCGGTGTCGATGGAATACCCCGCACCTCTATGCGGACCCCGC 1179
QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db 1180 TTCCGAGGTATACCCCTGCCCCAGGGCACGGAGTCTTCCGCTCTCTGGCTCCATCTG 1239
QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1240 CATGACCCCAACATCTTCAGGACCCAGAGAGTTCAACCCAGACCGTTTCTGTGATGCA 1299
QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1300 GATGACCGTTCAGGAAGCATCAGGCGTTCCTGCCCTTCTCTTAGGAGAGCGTGTCTGC 1359
QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
Db 1360 CTTGAGAGGGGCTGGCAAAAGCGAGTGTCTTCTCTTCAACACCATCTCTACAGCC 1419
```

```
QY 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1420 TTTCTCCTGGAGAGCCGTGCGCGCGGACACCTTGAGCCTCAAGCCACACCTCAGTGC 1479
QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1480 CTTTTCACATTCGCCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCAC 1539
QY 501 ThrGlnThrArg 504
Db 1540 ACGCAGACGAGA 1551

RESULT 10
LOCUS AX876954 2513 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 11859 from Patent EP1074617.
ACCESSION AX876954
VERSION AX876954.1 GI:40031690
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: Ep 1074617-A 11859 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1.2513
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
56..1750
/notes="unnamed protein product"
/codon_start=1
/protein_id="CA889506.1"
/db_xref="GI:40031691"
/translation="MEATGTWALLALALLLLTALSGTRARGHLPGGTPLPLGN
LIQRRGALYGLMRUSKIGPVFTIYLGPWRVVLVQEAVERALGGQAEFSGRG
TVAMLEGTFDHGIVFNSNGERKFTMLALRDLGMKREGEELIQAARCLVETP
QGTGEPFPPSLLAQTASNVVCSLLIFGRFVDEKFOAVRAAGTLLGVSSQSGQ
TYEMFSWFLRPLPGPHKQLLHVSTLAFTVRQVQHQGNLDASGPARDLVADPLKM
AOEONPQTEFTNKMMLTVIYLLPAGTWTSTVTCGYTLLLLMKYPHVQKVRBELNR
ELGGAQPSLGRDRLPYDVLVHEAQRLLALVPMGIPTLMKRTTRFRGTYLLPQGTVEV
FPLLSILHDPNI FKHPSEFNDFDLADGRFRKHEAFILFSLGKRVCLGEGIAKEL
FLFTTILQASLESPPDPTLSLKTPTVSGLFNIPPASTVGMDRVNVSVYTAGSHY
TPAVVRSLSHGPHALHTAAKWHNRTPIHNYKGHKATAGLAFHHRHKYSPSATTSI"

ORIGIN
Alignment Scores:
Pred. No.: 6,2e-194 Length: 2513
Score: 2527.00 Matches: 487
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.63% Indels: 0
DB: 6 Gaps: 0

US-10-669-693-2 (1-504) x AX876954 (1-2513)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuThr 20
Db 56 ATGAGGCGACCGGACCTGGCGCTGTCTGGCGCTGGCGCTCTCTCTCTCTGCTGAGC 115
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40
Db 116 CTGGCGCTGTCCGGACCGAGGCCCGACCTGCCCGCGGCCCGCCCGCGGTACCA 175
QY 41 LeuLeuGlyAsnLeuLeuLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
Db 176 CTGTGGGAAACCTCTCTGACGTACGGCCCGGGCGCTGTATTTCAGGCTCATGCGGCTG 235
```

QY 61 SerIysLysTyrGlyProValPheThrIleTyrLeuGlyProTyrArgProValValVal 80
 Db 236 AGTAAGAAGTACGACCGGTGTTCCACATCTACCTGGAGACCTGGCGGCTGTGGTGGTC 295
 QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
 Db 296 CTGGTGGGAGGAGGCTGTGGGAGGAGGCTGGGAGGTCAGGCTGAGGAGTTTCAGCGGC 355
 QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
 Db 356 CGGGAAACCGTAGCGATGCTGGAAGGAGCTTTGATGGCCATGGGGTTTCTTCTCCAAAC 415
 QY 121 GlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
 Db 416 GGGAGCGGTGGAGGAGCTGAGGAGTTTACCATGCTTGTCTGCGGAGCTGGGCATG 475
 QY 141 GlyLysArgGluGlyGluGluLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160
 Db 476 GGAAGCGGAGGAGGAGCTGATCCAGGCGGAGGCGCGGTGTCTGGTGGAGCATTC 535
 QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
 Db 536 CAGGGACAGAGGAGCGCCATTCGATCCCTCCCTGCTGCTGGCCAGGCGACCTCCAAAC 595
 QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
 Db 596 GTAGTCTGCTCCTCTCTTGGCTCGCTTCTCTATGAGATAAGGAGTTCCAGGCG 655
 QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr 220
 Db 656 GTGGTCCGGGAGCTGGTGGTACCTCTGCTGGAGTCTGCTCCAGGGGGTCTGACCTTAC 715
 QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
 Db 716 GAGATGTTCTCTCTGGTCTCTGGGCGCTTCCAGGCGCCCTCCAGGAGCTCTCCACAC 775
 QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
 Db 776 GTCCAGACCTTGGCTGCTTACAGTCCGCGAGTGCAGCAGCAGCAGGAGACCTGGAT 835
 QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
 Db 836 GCTTCCGGGCGCCAGCTGACCTTGTGATGCTTCTGCTGAAGATGGCACAGGAGAA 895
 QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
 Db 896 CAAAACCCAGGACAGATTCACCAACAGAACATGCTGATGACAGTCAITTTATTGCTG 955
 QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
 Db 956 TTTGCTGGGACGATGCGGTGAGCAGCAGCAGCAGCTGCTATACCTCTCTCTCTGATGAAA 1015
 QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
 Db 1016 TACCCTCATGTCAAAGTGGTACGAGAGCTGAATCGGAGCTGGGGGTGGCCAG 1075
 QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
 Db 1076 GCACCAAGCTTGGGACCGTACCCGCTCCCTTACACCGCAGCGGCTTCTGCTGATGGCG 1135
 QY 361 GlnArgLeuLeuAlaLeuValProMetGlyLysPheProArgThrLeuMetArgThrArg 380
 Db 1136 CAGCGGCTGTGGGCTGGTGGCCATGGAATATACCCGCACTCTATGCGGACACCCCGC 1195
 QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
 Db 1196 TTCGAGGGTACACCTGGCCCGGAGGAGTCTTCCCTTCTTCTTGGCTTCCATCTCTG 1255
 QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
 Db 1256 CATGACCCCAACATCTTCAAGCAGCAGGAGTTCAACCCAGAGCTTCCCTGGATGCA 1315

421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
 Db 1316 GATGAGCGTTTCAGAGCATGAGCGTCTCTGCCCTTCTCTTAGGAAAGCGTGTCTGC 1375
 QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPhePheThrIleLeuGlnAla 460
 Db 1376 CTTGAGAGGCGCTGGCAAAAGCGAGCTCTCTCTTCTTCACACCATCTCAAGCC 1435
 QY 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
 Db 1436 TTCTCCCTGGAGAGCGCGTGGCCCGGACACCTGAGCCTCAAGCCACCGTCAGTGCC 1495
 QY 481 LeuPheAsnIleProProAla 487
 Db 1496 CTTTTCACATTCCTCCCGAGCC 1516

RESULT 11
 BD156406
 LOCUS 2513 bp DNA linear PAT 17-JAN-2003
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD156406
 VERSION BD156406.1 GI:27862164
 KEYWORDS JP 2002191363-A/11249.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2513)
 AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 TITLE Primer for synthesizing full-length cDNA and use thereof
 JOURNAL Patent: JP 2002191363-A 11249 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 COMMENT OS Homo sapiens (human)
 PN JP 2002191363-A/11249
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C12P1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 PC, C12P21/02, C12P1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH key
 Location/Qualifiers
 (56)..(1747).
 FT CDS
 FEATURES
 source
 1..2513
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores: 6.2e-194 Length: 2513
 Pred. No.: 2527.00 Matches: 487
 Score: 2527.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 96.63%
 Indels: 0
 Gaps: 0
 DB: 6
 US-10-669-693-2 (1-504) x BD156406 (1-2513)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeu 20
 Db 56 ATGAGGCGGACCGGACCTGGCGCTGCTGTGGCGCTGGCGCTGCTCTCTGCTGCTGACG 115
 QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40
 Db 116 CTGGCGCTGTCCGGACCCAGGCGCCGAGCCACCTGCCCCCGGGGCCACCGCTACCA 175

QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
DB 176 CTGCTGGAAACCTCTCTCAGCTACGGCCCGGGCGCTGTATTACAGGCTCATGGCGGTG 235
QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrrArgProValVal 80
DB 236 AGTAAGAAGTACGACCGGTGTTCACCATCTACCTGGGACCTGGCGGCTGTGGTGGTC 295
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
DB 296 CTGGTTGGCAGAGGCTGTGGGAGGCCCTGGAGGTCTAGGCTGAGAGTTACGGCC 355
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
DB 356 CGGGAAACCGTAGCATGCTGGAAGGACTTTTTCATGSCCATGGGTTCCTTCTCCAAC 415
QY 121 GlyGluArgTrrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
DB 416 GGGAGCGGTGGAGGACGTGAGGAAGTTTACCATGCTTGTCTGCGGACCTGGGCATG 475
QY 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
DB 476 GGGAGCGAGAGAGCGAGGAGCTGATCCAGCGGAGGCCCGGTGTCTGGTGGAGACATTC 535
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
DB 536 CAGGGACAGAAAGACGCCCATTCGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAC 595
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
DB 596 GTAGTGTCTCCCTCCCTCTTGGGCTCCGCTTCTCTATGAGGATAAGGAGTTCCAGGCC 655
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerGlnGlyGlyGlnThrTyr 220
DB 656 GTGGTCCGGGACGTGTGTGTACCTGCTGGAGTACGCTCCCAAGGCGGTGACACCTAC 715
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240
DB 716 GAGATGTCTCTCTGTTCTGCGGCCCTGCCAGGCCCCCAAGCAGCTCTCCACAC 775
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
DB 776 GTCAGACACCTTGCTGCTCTTACAGTCCGGCAGTGCAGCAGCAGCAGGGAACCTGGAT 835
QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280
DB 836 GCTTCGGGCCCGCCAGCTGACCTTGTCATGCTGCTGCTGAGTGGCAGCAGGAGGAA 895
QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
DB 896 CAAACCCAGGCACAGAAATTCACCAACAGAACATGCTGATGACAGTCATTTATTTGCTG 955
QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
DB 956 TTTGCTGGGACGATGACGTCAGCACACCGTCCGCTGCTATACCTCTCTCTGTATGATA 1015
QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
DB 1016 TACCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGTGGCCAG 1075
QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
DB 1076 GCACCAAGCCTAGGGACCGTACCCGCTCCCTTACACCGACCGGTCTGCGATGAGGCG 1135
QY 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
DB 1136 CAGCGGCTGTGGCTGTGGTGGCCATGGGAATACCCCGCACCTCATGCGGACCCCGC 1195
QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
DB 1196 TTCCCGGGGTACACCTGCCCGAGGACGAGGAGTCTTCCCTCTCTGCTCCATCTCTG 1255
QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420

DB 1256 CATGACCCCAACATCTTCAAGACCCAGAGAGTTCAACCCAGACCGTTTCTGTGATGCA 1315
QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
DB 1316 GATGACCGTTTCAGGACGATGAGCGTCTCTGCGCTTCTCTTAGGAAGCGTGTCTGTC 1375
QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPhePheThrIleLeuGlnAla 460
DB 1376 CTTGGAGAGGCGCTGGCAAAAGCGAGCTCTCTCTCTTCCACCCATCTCAAGGCC 1435
QY 461 PheSerLeuGluSerProCysProAspProPheLeuSerLeuLysProThrValSerGly 480
DB 1436 TTTCTCCCTGGAGAGCCCGTCCCGGACACCTTGAGCTCAAGCCACCGTCAGTGGC 1495
QY 481 LeuPheAsnIleProProAla 487
DB 1496 CTTTTCACATTTCCCCCAGCC 1516
RESULT 12
AK027605
LOCUS
DEFINITION Homo sapiens cDNA FLJ14699 fis, clone NT2RP2006571, moderately
similar to CYTOCHROME P450 2G1 (BC 1.14.14.1).
ACCESSION AK027605
VERSION AK027605.1 GI:14042395
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS
1 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Ishibashi, T., Fujimori, K.,
Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Ishii, S., Kawai, Y.,
Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K.,
Masuho, Y. and Kanehori, K.
NEDO human cDNA sequencing project
Unpublished
REFERENCE
JOURNAL 2 (bases 1 to 2513)
AUTHORS Isogai, T. and Otsuki, T.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
1..2513
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2006571"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
56..1750
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:14042396"
/translation="MEATGTWALLALALLLTALSGTRAGHLPGCTPLILGN
LLQLRPGALYGLMRLSKYGPVFTIYLPWPVPVVLVQAEAVRALGQAEPSGRG
TVAMLEGTFDGHGVFFSNGERWRQRKRTMLALRDLGMKREGELIQAEARCLVETFG
QGTGRPPDPSLLLAQAQTSNVVCSLLFLGLRFSYEDKRFQAVVRAAGTLLGVSSQGG

TYMFNPLRLPQPHKQLLHVSTLAAFTVRVQVQHQGNLDASGPARDLVDAFLKX
AEQNFCTETNQMILTVIYLLFAGTMTVSTVYGLLLMLKYPHVQKVRBELNR
ELGAGQAPSLGDRTRLYTDAVLHEAQRLLALVPMGIPTRIMTRFRGYTLPGQTEV
PFLGSLIHDNPIFKHPEEPNDRFLDAGRFRKHEAPLPSLGRVCLGELGAKAEL
FLFTTILQAPSLSPCPDPLSLKPTVSGLEPNIPASTVGMDRVNVSRVYTAGSHY
TPAVVFRSLSHGPHLTHAAKMHNRPIPHNYKGHKATAGLA FHRHKVSPSAITST"

ORIGIN

Alignment Scores:
Pred. No.: 6,2e-194 Length: 2513
Score: 2527.00 Matches: 487
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.63% Indels: 0
DB: 9 Gaps: 0

US-10-669-693-2 (1-504) x AK027605 (1-2513)

Qy	1	MetGluAlaThrGlyThrTrrPAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr	20
Db	56	ATGGAGGCGACCGGCACCTGGCGCTGCTGGCGCTGGCGCTGCTGCTGCTGCTGAGC	115
Qy	21	LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro	40
Db	116	CTGGCGCTGCTCCGGACCCAGGCGCCGAGGCGACCTGCCCGCGGCGCCAGCGCTACCA	175
Qy	41	LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu	60
Db	176	CTGCTGGGAAACCTCTCGCAGCTACGGCGCGGCGCTGTATTACGGGCTCATGGCGGTG	235
Qy	61	SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrrPArgProValValVal	80
Db	236	AGTAAGAAGTACGGACCGGTGTACCACTACCTGGGACCTCTGGCGGCTGTGGTGGTC	295
Qy	81	LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly	100
Db	296	CTGGTTGGCAGAGGCTGTGGGAGGCGCTGGGAGGTTCAGGCTGAGGAGTTCAGCGGC	355
Qy	101	ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn	120
Db	356	CGGGAAACCGTAGCGATGCTGGAAGGGACTTTTGATGGCCATGGGGTTTTCTTCTCCAAC	415
Qy	121	GlyGluArgTrrPArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet	140
Db	416	GGGAGCGGTGGAGGCGAGCTGAGGAAGTTTACCATGCTTGCTGCGGGACCTGGGCATG	475
Qy	141	GlyLysArgGluGlyGluGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe	160
Db	476	GGGAACGGAAGCGAGGAGCTGATCCAGCGGAGGCGCGGTGCTGCTGGGAGACATTC	535
Qy	161	GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn	180
Db	536	CAGGGACAGAGGACGCCATTCGATCGCTCCCTGCTGCTGGCGCCAGGCGCACCTCCAAC	595
Qy	181	ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla	200
Db	596	GTAGTCTGCTCCCTCTCTTTGGCCCTCCGCTTCTCTATGAGGATGAAGGTTCCAGGCC	655
Qy	201	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
Db	656	GTGTCTCGGCGACCTGGTGTACCTGCTGGGAGTCAGCTCCCGAGGGGTTCAGACCTAC	715
Qy	221	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis	240
Db	716	GAGATGTTCTCTGTTCTGCGCGCCCTGCCAGGCGCCCGCACAGAGCTCTCCACCCAC	775
Qy	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
Db	776	GTGAGCACCTTGCTGCTCTTACAGTCCGCGAGGTGCAGCAGCAGCAGGCGAATCGGAT	835
Qy	261	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu	280
Db	836	GCTTCGGGGCCCGCACGTGACCTTGTGATGCTCTTCTGCTGAAGATGGCACAGGAGAA	895

Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
Db	896	CAAAACCCAGGACAGAAATTCACCAACAGAACATGCTGATGACAGATCAATTAATTGCTG	955
Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys	320
Db	956	TTTGCTGGGAGATGATGCGTCCAGCACCGGTGCGGTATACCTCTCTCTGATGAAA	1015
Qy	321	TyrProHisValGlnLysTrrPValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln	340
Db	1016	TACCTCTATGTCCTCAAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAG	1075
Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTrrThrAspAlaValLeuHisGluAla	360
Db	1076	GCACCAAGCCTAGGGGACCGTACCGGCTCTCTTACACCGACCGCGGTCTTGATGAGGGG	1135
Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
Db	1136	CAGCGGCTGCTGGCGCTGGTCCCATGGGAATACCCCGCACCTCATGGGACCCCGC	1195
Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1196	TTCCGAGGATACACCTGCGCCAGGCGACGGAGGTCTTCCCGCTCTTGCTCCATCTCTG	1255
Qy	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1256	CATGACCCCAACATCTTCAGCACCCAGAGAGTTCAACCCAGACCGTTTCTGGATGCA	1315
Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1316	GATGACCGGTTTCAGGAAGCATGAGGCGTTCCTCGCGCTTCTCTAGGGAAGGTGTCTGC	1375
Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAla	460
Db	1376	CTTGAGAGGGCTCGGCAAAAGCGAGGCTTTCCTCTCTTCCACCACTCTTACAGCC	1435
Qy	461	PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1436	TTTCTCTCGAGAGCGGCTGCGCGCGGACACCTGAGCCTCAAGCCACCGCTCAGTGCC	1495
Qy	481	LeuPheAsnIleProProAla	487
Db	1496	CTTTTCAACATTCCTCCCGACGCC	1516

RESULT 13

AX552231
LOCUS
Sequence 140 from Patent WO0162927.
DEFINITION
AX552231
ACCESSION
AX552231.1 GI:25896469
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Barville, S.C., Greenawalt, L.B., Lincoln, S.E., Stockdreher, T.K.,
Amshley, S., Chang, S.C., Chen, W., D'Sa, S.A., Dam, T.C., Liu, T.F.,
Rosen, B.H., Russo, F.D., Spiro, P.A., Bradley, D.L., Chen, A.,
Cohen, H.J., Daffo, A., Daniels, S.E., Dufour, G.E., Flores, V.,
Fong, W.T., Hodgson, D.M., Jackson, S., Jones, A.L., Panzer, S.,
Roseberry, A.M., Shah, P., Wright, R.J., Yap, P.E., Yu, J.Y.,
Bratcher, S.R., Chalup, M.S., Dahl, C.R. and Hillman, J.L.
Polypeptides and corresponding polynucleotides for diagnostics and
therapeutics
Patent: WO 0162927-A 140 30-AUG-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..2114
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

MEDLINE	93349061
PUBMED	8346920
REMARK	GenBank

GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 135806] from the original journal article. This sequence comes from Table I.

FEATURES

source

```

source
1. .2028
/organism="Oryctolagus cuniculus"
/mol_type="mRNA"
/db_xref="taxon:9986"
1. .2028
/genes="cytochrome P450 2B-Bx"
1. .1476
/genes="cytochrome P450 2B-Bx"
/note="phenobarbital-inducible; This sequence comes from
Table 1"
/codon_start=1
/product="cytochrome P450 2B-Bx"
/protein_id="AAB27705.1"
/db_xref="GI:402843"
/translation="MEFSLLLALLFLAGLLLLFRGHKPAHGRLPGGSPILPVLGNL
QMDKGLLRFLRLREKYGVDFVTLVGRSPVVLGGTDAIEALVDOAAEAFSGRGKI
VWDFIPQGVGVFANGERWALRLRESLATMDYDFMGKGRSVEERIQEARCGLVELRL
KGAIDLNTLLFHSITLNCISVIPGKRDYKDPVFLRLDUFFQPSFLISQSVQKVE
LPSGELKHPGPTHQRLYRNQLQINTFTTQSVKEKRATLDPSNPDFIDVLLRMKKE
SPSFEFHHQNLILVLAUFAGTETTTTLRYGLFLMLKYPHVTERVQKEIEQVIG
HAPPAIDLKRMQYPTDAVHIEHQRGLDILIPGVPHVTKTQTFQGVYIPKNTVEFV
SALHDDRYFKEFTFNGFLADNALKRNEGMPFSLGKRVCLGEGIARTLELFLF
TTILQNFSTIASVPFEDIDLTPRESGVNGVPPSYOIRFLAP"

```

ORIGIN

Alignment Scores:

Pred. No.:	3.16-96	length:	2028
Score:	1313.00	Matches:	247
Percent Similarity:	71.05%	Conservative:	99
Best Local Similarity:	50.72%	Mismatches:	137
Query Match:	50.21%	Indels:	4
DB:	4	Gaps:	3

US-10-669-693-2 (1-504) x S64259 (1-2028)

QY		7	TpAlaLeuLeuLeuAlaLeuAlaLeuLeu-----LeuLeuLeuThrLeuAlaLeuSer	24
Dd		7	TTACGCTGCCTCCTCCTCGCTTCTTCGCAGGGCCCTCGTGCTCTCTGTCAGGGGC	66
QY		25	GlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuProLeuGlyAsn	44
Dd		67	CACCCAGGGCCACGGCGGCTCCCCAGAGCCCTCCCTCTGCCCGTCTCTGGGGAAC	126
QY		45	LeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerLysIstyr	64
Dd		127	CTTCTGCAGATGCAGAGAAGGGCTCTCGCTCGCTCTCTCGGGCTCCGAGAAATAC	186
QY		65	GlyProValPheThrlleTyLeuGlyProTrpArgProValValLeuValGlyGln	84
Dd		187	GGGACGTGTTCACGGTGTACTCGGATCC---AGACCCTGTGTGTCTGTCTGGGACG	243
QY		85	GluAlaValArgGluAlaLeuGlyGlyGlnAlaGluPheSerGlyArgGlyThrVal	104
Dd		244	GATGCCATCCGCGAGGCCCTCGTGGACCAAGCTGAGGCCCTTTCTGCGAGGGGGAAGTC	303
QY		105	AlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTrp	124
Dd		304	GCCGTGTGGATCGATCTTCCAGGGATACGGAGTGATCTTTCGCCAACGGGAGCGCTGG	363
QY		125	ArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuClyMetGlyLysArgGlu	144
Dd		364	CGGGCCCTTCGAGATTCTCCCTTGGCCACCATCGCGAGCTTCGGCATGGGGAGCGGAGC	423
QY		145	GlyGluGluLeuIleGlnAlaGluAlaAraCysLeuValGluThrPheGlnGlyThrGlu	164
Dd		424	GTGAGAGAGCGCATTCAGGAGGAGGCCGCTGTCTCTGTGTGAGAGCGTGGGAAAAC	483

Search completed: September 18, 2004, 03:29:08
Job time : 5903 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 00:56:16 ; Search time 123 Seconds
(without alignments)
2273.945 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLLT.....PPAFQVQRPTDLHSTQTR 504

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10669693/runat_15092004_101311_5876/app.query.fasta_1.647
-DB=Issued Patents NA -QWMT=fastcap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10669693 @CGN 1.1.69 @runat_15092004_101311_5876 -NCPU=6 -ICPU=3
-NO.MWAP -LARGSQURY -NEG SCORES=0 -WAIT -DSBLOK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1238	47.3	1476	4	US-08-277-031B-10
2	1232	47.1	2907	4	US-09-023-655-1053
3	1213.5	46.4	1485	4	US-08-277-031B-8
4	1213.5	46.4	1485	4	US-08-277-031B-9
5	1185	45.3	1737	2	US-08-750-703-2
6	1175.5	45.0	1740	2	US-08-750-703-1
7	1108	42.4	1825	4	US-09-023-655-1061
8	1087.5	41.6	1854	1	US-08-201-118-4
9	1087.5	41.6	1854	2	US-08-238-821B-4
10	1087.5	41.6	1854	4	US-09-023-655-1056
11	1087.5	41.6	1854	4	PCT-US95-05744-4
12	1086.5	41.5	1473	4	US-08-277-031B-2

13	1085.5	41.5	1473	4	US-08-277-031B-15
14	1085.5	41.5	1845	4	US-09-023-655-1057
15	1085.5	41.5	1852	1	US-08-201-118-10
16	1085.5	41.5	1852	2	US-08-238-821B-10
17	1085.5	41.5	1852	5	PCT-US95-05744-10
18	1084.5	41.5	1746	1	US-08-201-118-2
19	1084.5	41.5	1746	2	US-08-238-821B-2
20	1084.5	41.5	1746	4	US-09-023-655-1059
21	1084.5	41.5	1746	5	PCT-US95-05744-2
22	1078.5	41.2	1591	2	US-08-194-981E-3
23	1073.5	41.1	1473	4	US-08-277-031B-14
24	1073.5	41.1	1995	4	US-09-023-655-1055
25	1073.5	41.1	2009	1	US-08-201-118-6
26	1073.5	41.1	2009	5	US-08-238-821B-6
27	1073.5	41.1	2009	5	PCT-US95-05744-6
28	1067.5	40.8	2258	1	US-08-201-118-12
29	1067.5	40.8	2258	2	US-08-238-821B-12
30	1067.5	40.8	2258	4	US-09-023-655-995
31	1067.5	40.8	2258	5	PCT-US95-05744-12
32	1059	40.5	1419	2	US-08-194-981E-4
33	1054.5	40.3	1473	4	US-08-277-031B-11
34	1051.5	40.2	1473	4	US-08-277-031B-12
35	1050.5	40.2	1829	1	US-08-201-118-8
36	1050.5	40.2	1829	5	US-08-238-821B-8
37	1050.5	40.2	1829	5	PCT-US95-05744-8
38	1044.5	39.9	1473	4	US-08-277-031B-13
39	1033.5	39.5	1892	1	US-08-201-118-14
40	1033.5	39.5	1892	2	US-08-238-821B-14
41	1033.5	39.5	1892	5	PCT-US95-05744-14
42	982.5	37.6	1482	2	US-08-194-981E-1
43	982.5	37.6	1482	4	US-08-277-031B-3
44	960.5	36.7	1419	2	US-08-194-981E-2
45	871.5	33.3	1857	4	US-09-023-655-902

ALIGNMENTS

RESULT 1

US-08-277-031B-10
; Sequence 10, Application US/08277031B
; Patent No. 6620593
; GENERAL INFORMATION:
; APPLICANT: Hayashi, Koji
; APPLICANT: Sakaki, Toshiyuki
; APPLICANT: Yabusaki, Yoshiyasu
; APPLICANT: Komai, Koichiro
; APPLICANT: Kaneko, Hideo
; APPLICANT: Nakatsuka, Iwao
; TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF
; TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING
; TITLE OF INVENTION: HUMAN CYTOCHROME P450
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5inch, 1.44MB
; COMPUTER: IBM PC
; OPERATING SYSTEM: Dos 5.0
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/277,031B
; FILING DATE: 19-JULY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP-201120/1993
; APPLICATION NUMBER: JP-180246/1993
; APPLICATION NUMBER: JP-208279/1993
; FILING DATE: 20-07-1993


```

; FILING DATE: 21-07-1993
; FILING DATE: 30-07-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Raymond C. Stewart
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 20-3530P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-277-031B-10

Alignment Scores:
Pred. No.: 3,39e-125 Length: 1476
Score: 1238.00 Matches: 236
Percent Similarity: 68.51% Conservative: 97
Best Local Similarity: 48.67% Mismatches: 150
Query Match: 47.34% Indels: 4
DB: 4 Gaps: 3

US-10-669-693-2 (1-504) x US-08-277-031B-10 (1-1476)
QY 8 AlaLeuLeuAlaLeuAlaLeu-----LeuLeuLeuThrLeuAlaLeuSerGly 25
Db ::::|||||
Dy AGCGTCCTCCTTCTCTTGCACTCCTCACAGGACTCTTGCTACTCTGTTCAGCGCCAC 69
QY 26 ThrArgAlaArgGlyHisLeuProProGlyProThrProLeuProLeuLeuGlyAsnLeu 45
Db ::::|||||
Dy CCTAACACCATGACCGCCTCCCACAGGGCCCCGCCCTCTGCCTCTTTTGGGAACCTT 129
QY 46 LeuGlnLeuArgProGlyAlaLeuTyrrSerGlyLeuMetArgLeuSerLysIstyrGly 65
Db ::::|||||
Dy CTGCAGATGGATAGAGAGGGCTTCTCAAATCCTTTCTGAGGTTCCGAGAGAAATATGGG 189
QY 66 ProValPheThrIleTyrLeuGlyProTrpArgProValValValLeuValcylGlnGlu 85
Db ::::|||||
Dy GACGTCCTCAGGTACACCTGGGACCG--AGGCCCTGGTTCATGCTGTGTGGAGTAGAG 246
QY 86 AlaValArgGlnAlaLeuGlyGlyGlnAlaGluGluPheSerGlyArgGlyThrValAla 105
Db ::::|||||
Dy GCCATACGGAGGCCCTGTGACAAGCTGAGGAGCTGAGGCTTCTCTGGCGGGAAAAATCGCC 306
QY 106 MetLeuGluGlyThrPheAspGlyHisIsglyValPhePheSerAsnGlyGluArgTrpArg 125
Db ::::|||||
Dy ATGGTCACCCATTCTTCGGGGGATATGTTGATCTTTGGCAATGGAAACCGCTGGGAAG 366
QY 126 GlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLysArgGluGly 145
Db ::::|||||
Dy GTGCTTCGGCATTCCTGTGACCACTATGAGGACTTCGGGATGGGAAGCGGAGTGTG 426
QY 146 GluGlnLeuIleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGly 165
Db ::::|||||
Dy GAGGAGCGGATTACAGGAGGAGCTCAGTGTCTGTATAGAGAGCTTCGGAATATCCAAGGGG 486
QY 166 ArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeu 185
Db ::::|||||
Dy GCCTTCATGGAGCCCACTTCCTTCCAGTCCATTCAGCCCAACATCATGTCTCCATC 546
QY 186 LeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAlaValValArgAlaAla 205
Db ::::|||||
Dy GTCTTTGGAAAACGATTCCACTACCAAGATCAAGAGTCTCTGNAGATGCTGAACTGTTC 606
QY 206 GlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrThrGluMetPheSerTrp 225
Db ::::|||||
Dy TACCAGACTTTTTCATCATCAGCTCTGTATCCGGCCAGCTGTTTCAGTCTTCTCTGGC 666
QY 226 PheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHisValSerThrLeuAla 245

```

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/023,655
 FILING DATE: HEREWITH
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Zeller, Karen J.
 REGISTRATION NUMBER: 37,071
 REFERENCE/DOCKET NUMBER: PA-0001 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166
 INFORMATION FOR SEQ ID NO: 1053:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2907 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GENBANK
 CLONE: g181293
 US-09-023-655-1053

Alignment Scores:
 Pred. No.: 4.52e-124 Length: 2907
 Score: 1232.00 Matches: 235
 Percent Similarity: 68.71% Conservative: 101
 Best Local Similarity: 48.06% Mismatches: 149
 Query Match: 47.11% Indels: 4
 DB: 4 Gaps: 3

US-10-669-693-2 (1-504) x US-09-023-655-1053 (1-2907)

```

102 166 ArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeu 185
103 493 GGCCTCGTGGACCCACCTTCTTCCATTACCGCCAAATCATCTCTCCATC 552
104 186 LeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAlaValValArgAlaA 205
105 553 ATCTTTGGAAACGCTTCCACTACCAAGATCAAGAGTTCTCTGAAGACGCTGAATTTGTT 612
106 206 GlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyrGluMetPheSerT 225
107 613 TGCCAGAGTTTCTTACTCATCAGCTTATATCCAGCCAGCTGTTGAGCTCTTCTCTGGC 672
108 226 PheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisValSerThrLeuAla 245
109 673 TTCTTGAATACTTCTCTGGGCACACAGCAAGTTTACAAAAACCTCAGGAAATCAAT 732
110 246 AlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSerGlyProAla 265
111 733 GCTTACATTGGCCACAGTGTGGAGAGCAGCCGTGAACCCCTGGACCCCGCCGCC--- 789
112 266 ArgAspLeuValAlaPheLeuLeuLysMetAlaGlnGluGlnAlaAsnProGlyThr 285
113 790 AGGACCTCATCGACACCTACTCTCCACATGGGAAAGAGAAATCCACCCACACAGT 849
114 286 GluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAlaGlyThrMet 305
115 850 GAATTCAGCCACCAAGACCTCATCAACACGCTCTCGCTCTCTTTGCTGGCACTGAG 909
116 306 ThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLysTyrProHisValGln 325
117 910 ACCACGACACACTCTCCGCTACGGCTTCTGTCTCATGCTCAAAATACCTCATGCGCA 969
118 326 LysTrpValArgGluLeuAsnArgGluLeuGlyAlaGlyGlnAlaProSerLeuGly 345
119 970 GAGAGAGTCTACAAGGAGATTGAACAGGTGTGGCCACATCGCCCTCCAGCGCTTGAT 1029
120 346 AspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAlaGlnArgLeuAla 365
121 1030 GACGAGCCAAATGCCATACACAGAGCAGTCATCCGTGAGATTCAGAGATTTGCTGAC 1089
122 366 LeuValProMetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTyrThr 385
123 1090 CTTCTCCCATGGGTGGCCCAATTTGTCTACCCCAACACACACAGCTTCTGAGGGTACACC 1149
124 386 LeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIle 405
125 1150 ATCCCCAAGGACACGGAAGTATTCTCATCTCTGAGCAGCTGCTCTCCGTGACCCACACTAC 1209
126 406 PheLysHisProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArg 425
127 1210 TTTGAAACACGACGCTTCAATCTCTGACCACTTTCTGGATGCCAATGGGCACTGAAA 1269
128 426 LysHisGluAlaPheLeuProPheSerLeuLysArgValCysLeuGlyGluGlyLeu 445
129 1270 AGAATGAAGCTTTTATCCCTTCTCTAGGGAAGCGGATTTGTCTTGTGTAAGGCACT 1329
130 446 AlaLysAlaGluLeuPhePheThrThrIleLeuGlnAlaPheSerLeuGluSer 465
131 1330 GCCCGTGGCAATTGTTCTCTTCTTCCACACCATCTCTCCAGAACTTCTCCGTGGCCAGC 1389
132 466 ProCysProProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIlePro 485
133 1390 CCGTGGCTCTGAAGACATCGATCTGACACCCCGAGGAGTGTGGTGGGCAAAATACCC 1449
134 486 ProAlaPheGlnLeuGlnValArgPro 494
135 1450 CCAACATACAGATCTGCTTCTCTGCCC 1476
136
137 RESULT 3
138 US-08-277-031B-8
139 ; Sequence 8, Application US/08277031B
140 ; Patent No. 6620593
141 ; GENERAL INFORMATION:

```

[illegible]

[illegible]

Db 682 AAACACCTGCCAGGACACACAGAGCCCTTTCAGTTGCTGCAAGGCTCGAGGACTTC 741
Qy 248 ThrValArgGlnValGlnGlnHscGlnGlyAenLeuAspAlaSerGlyProAlaArgAsp 267
Db 742 ATAGCAAGAGGTGGACACACAGGCGACGCTGGATCCCAATTCGCCA---CGGGAC 798
Qy 268 LeuValAspAlaPheLeuLeuLeuMetAlaGlnGlnGlnAsnProGlyThrGluPhe 287
Db 799 TTCATTGACTCTCTTCATCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 858
Qy 288 ThrAsnLysAsnMetLeuMetThrValIleThrLeuLeuPheAlaGlyThrMetThrVal 307
Db 859 TACTTGAAGAACTGATGATGAGCAGCTTGAACCTTTCATTCAGGACACCGAGCGTC 918
Qy 308 SerThrThrValGlyThrLeuLeuLeuMetLysThrProHisValGlnLysThr 327
Db 919 AGCACCACCTGACATGCTTCTGCTGCTCATGAAGCACCAGAGGTGGAGGCCAAG 978
Qy 328 ValArgGlnGlnLeuAsnArgGlnLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArg 347
Db 979 GTCCATGAGGAGATTCACAGATGATCGCAAGACCGGACGCCCAAGTTTGAGGACCG 1038
Qy 348 ThrArgLeuProThrThrAspAlaValLeuHisGlnAlaGlnArgLeuLeuAlaVal 367
Db 1039 GCCAAGATCCCTACATGAGGAGGAGTATCCAGAGATCCAAAGATTTCGAGACGTGATC 1098
Qy 368 ProMetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyThrThrLeuPro 387
Db 1099 CCATGAGTTGGCCGCGAGAGTCAAAAGAGGACACCAAGTTTCGGGATTTCTTCTCCCT 1158
Qy 388 GlnGlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePheLys 407
Db 1159 AAGGGCATGAAGTGTCCCTATGTTGGGCTCCGTCGTAGAGACCTCAGGTCTCTCC 1218
Qy 408 HisProGluGlnPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHis 427
Db 1219 AACCCCGGACCTTCAATCCCGACCTTCTCGGTGAGAGGGGCGAGTTTAAGAAGCGT 1278
Qy 428 GluAlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGlnGlyLeuAlaLys 447
Db 1279 GATGCTTTTGTGCGCTTCTCCATCAGAAAGCGGAACCTGTTTCGGAAGGCGCTGGCCAGA 1338
Qy 448 AlaGluLeuPheLeuPheThrThrIleLeuGlnAlaPheSerLeuGluSerProCys 467
Db 1339 ATGAGCTCTTCTCTCTTCCACACCGTATCGAGAACTTCGCGCTCAGTCTCCAG 1398
Qy 468 ProProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAla 487
Db 1399 TCACCTAAGGACATGACGTGTCCTCCCAACACGTTGGGCTTTCGACGATCCACGAAAC 1458
Qy 488 PheGlnLeuGlnValArgPro 494
Db 1459 TACACCATGAGCTTCTGTCCTC 1479

RESULT 7

US-09-023-655-1061
Sequence 1061, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 5.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:
FILING DATE: US/09/023,655

PRIOR APPLICATION NUMBER: US/09/023,655

CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US

TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1061:
SEQUENCE CHARACTERISTICS:
LENGTH: 1825 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g181357

US-09-023-655-1061

Alignment Scores:
Pred. No.: 7,37e-111 Length: 1825

Score: 1108.00 Matches: 230

Percent Similarity: 64.73% Conservative: 93

Best Local Similarity: 46.09% Mismatches: 162

Query Match: 42.37% Indels: 14

DB: 4 Gaps: 6

US-10-669-693-2 (1-504) x US-09-023-655-1061 (1-1825)

Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuLeuLeuLeuLeuLeuLeu 19

Db 56 ATGACAGCATATAGCACAGCCATCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 115

Qy 20 ThrLeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeu 39

Db 116 ACCTTAGC-----TCAAGAGATAGGAGAAAGCTGCTCCGGGACCCAGACCCCTC 166

Qy 40 ProLeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArg 59

Db 167 TCAATCTGGGAACCTGCTGCTTTGCTCCCAAGACATGCTGACTCTCTCCTCACTAAG 226

Qy 60 LeuSerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValVal 79

Db 227 CTGACGAGGAGTATGGCTCCATGTACACAGTGCACCTGGGACCC---AGCGGGTGGTG 283

Qy 80 ValLeuValGlyGlnGlnAlaValArgGlnAlaLeuGlyGlyGlnAlaGluGluPheSer 99

Db 284 GTCTCTCAGCGGTACCAAGCTGTGAAGAGGCGCTGTGTGGACAGGAGGAGGAGTTAGT 343

Qy 100 GlyArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSer 119

Db 344 GGGCGCGGTGACTACCTGCTCTTTTCAACTTTTACCAAGGGCAATGCGATCGCTCTTCC 403

Qy 120 AsnGlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGly 139

Db 404 AGTGGGATCGATGGAGAGGCTCTGAGACAGTTCTCTATCCAGATTCTACGGAATTCGGG 463

Qy 140 MetGlyLysArgGluGlyGluGluLeuIleGlnAlaAlaArgCysLeuValGluThr 159

Db 464 ATGGGGAAGAGAGCATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 523

Qy 160 PheGlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSer 179

QY	29	ArgGlyHisLeuProGlyProThrProLeuProLeuGlyAsnLeuLeuGlnLeu	48
Db	88	AGAGGAAAACTCCCTCCTGCCTCCCTCTCCAGTGATTGGAATATCTCCACAGATA	147
QY	49	ArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerLysLysTyrGlyProValPhe	68
Db	148	GGTATTAAAGACATCAGCAAACTCTTAACCAATCTCTCAAGAGCTATGSCCTGTGTC	207
QY	69	ThrIleTyrLeuGlyProThrArgProValValLeuValGlyGlnGluAlaValArg	88
Db	208	ACTCTGTATTTGGC---CTGAAGCCATAGTGGTGTGCATGCATATGAAGCAGTGAAG	264
QY	89	GluAlaLeuGlyGlyGlnAlaGluGluPheSerGlyArgGlyThrValAlaMetLeuGlu	108
Db	265	GAAGCCCTGATTGATCTGGAGAGGAGTTTCTGGAAGAGGCATTTCCTCCAGCTGAA	324
QY	109	GlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTrpArgGlnLeuArg	128
Db	325	AGAGCTAACAGAGGATTTCCGAATTGTTTTCAGCAATGGAAGAAATGGAAGGAGATCCG	384
QY	129	LysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLysArgGluGlyGlnLeu	148
Db	385	CGTTTCTCCCTCATGACGCTGCGGAATTTGGGATGGGAAGAGGAGCATTCAGGACCGT	444
QY	149	IleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPhe	168
Db	445	GTTCAAGAGGAAGCCGCTGCCTGTGGAGGAGTTGCAGAAAAACCAAGCCCTCACCTGT	504
QY	169	AspProSerLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGly	188
Db	505	GATCCCATTTTCATCCTGGGCTGCTCCTCCCAATGTGATCGTCCATATTATTTCCAT	564
QY	189	LeuArgPheSerTyrGluAspLysGluPheGlnAlaValAlaArgAlaGlyGlyThr	208
Db	565	AAACGTTTTTGATTAAAGATCAGCAATTTCTTAACCTTAATGAAAGAGTGAATGAAAC	624
QY	209	LeuLeuGlyValSerSerGlnGlyGlyGlnThrTyrGluMetPheSerTrpPheLeuArg	228
Db	625	ATCAGATTTTGGACAGCCCTGGATCCAGATCTGCAATAATTTTCTCCTATCATTCAT	684
QY	229	ProLeuProGlyProHisLysGlnLeuLeuHisValSerThrLeuAlaAlaPheThr	248
Db	685	TACTTCCGGAACTCAACAAATTTACTTAAACCGTTGCTTTTATGAAAGTTATATT	744
QY	249	ValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeu	268
Db	745	TTGAAAAAGTAAAGAACACCAAGATCAATGACATGAACAACCTT---CAGACTTT	801
QY	269	ValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAsnProGlyThrGluPheThr	288
Db	802	ATTGATTGCTTCTGTGATAAATGGAGAGGAAAAAGCACAACCACTCTGAATTTACT	861
QY	289	AsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSer	308
Db	862	ATTGAAAGCTTGGAAAACTGCACTGTGACTTTGTTGGAGCTGGACAGACAGACAAGC	921
QY	309	ThrThrValGlyTyrThrLeuLeuLeuMetLysTyrProHisValGlnLysTrpVal	328
Db	922	ACAACCTGAGATATGCTCTCTCTCTGCTGAGCCCGAGAGGTCACAGCTTAAAGTC	981
QY	329	ArgGluGlnLeuAsnArgGluLeuLeuGlyAlaGlyAlaProSerLeuGlyAspArgThr	348
Db	982	CAGGAAGAGATTGAACCTGTGATTGGCAGAAACCGGAGCCCTGCATGCAAGACAGAGC	1041
QY	349	ArgLeuProTyrThrAspAlaValLeuHisGluAlaGlnArgLeuLeuAlaLeuValPro	368
Db	1042	CACATGCCCTTACACAGATGCTGTGGTGCACGAGTCCAGAGATACCTTGAACCTTCTCC	1101
QY	369	MetGlyIleProArgThrThrLeuMetArgThrThrArgPheArgGlyTyrThrLeuProGln	388
Db	1102	ACCAGCTGCCCATCGAGTGACCTGTGACATTAAATTCAGAACTATCTCATCTCCAAG	1161

QY	399	GlyThrGluValPheProLeuLeuGluGlySerIleLeuHisAspProAsnIlePheLysHis	408
Db	1162	GGCACCAACATATTAATTTCCCTGACTCTCTGTGTACATGACACAAAGAATTTCCCAAC	1221
QY	409	ProGluGluPheAsnProAspArgPheLeuAspAlaaspGlyArgPheArgLysHisGlu	428
Db	1222	CCAGAGATGTTTGACCCCTCATCACTTTCTGGATGAAGGTGGCAATTTTAAAGAAAGTAAA	1281
QY	429	AlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGlyLeuAlaLysAla	448
Db	1282	TACTTCATGCTTTCACAGCAGGAAACGAGATTGTGTGGAGAGACCCCTGGCCGGCATG	1341
QY	449	GluLeuPheLeuPhePheThrIleLeuGluAlaPheSerLeuGluSerProCysPro	468
Db	1342	GAGCTGTGTTTATTCCTGACCTCCATCTTTACAGAACTTTAAACCTGAAATCTCTGGTTGAC	1401
QY	469	ProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAlaPhe	488
Db	1402	CCAAAGAACCTTGACACCACCTCCAGTTGTCAATGGTTTGGCTCTGTGGCGCCTTCTAC	1461
QY	489	GlnLeu 490	
Db	1462	CAGCTG 1467	
RESULT 10			
US-09-023-655-1056			
; Sequence 1056, Application US/09023655			
; Patent No. 6607879			
; GENERAL INFORMATION:			
; APPLICANT: Cocks, Benjamin G.			
; APPLICANT: Susan G. Stuart			
; APPLICANT: Jeffrey J. Seilhamer			
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE			
; TITLE OF INVENTION: EXPRESSION			
; NUMBER OF SEQUENCES: 1508			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.			
; STREET: 3174 PORTER DRIVE			
; CITY: PALO ALTO			
; STATE: CALIFORNIA			
; COUNTRY: USA			
; ZIP: 94304			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/023,655			
; FILING DATE: HERewith			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; CLASSIFICATION:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Zeller, Karen J.			
; REGISTRATION NUMBER: 37,071			
; REFERENCE/DOCKET NUMBER: PA-0001 US			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (650) 855-0555			
; TELEFAX: (650) 845-4166			
; INFORMATION FOR SEQ ID NO: 1056:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1854 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; IMMEDIATE SOURCE:			
; LIBRARY: GENBANK			
; CLONE: g181301			
US-09-023-655-1056			

Alignment Scores:
Pred. No.: 1,31e-108 Length: 1854
Score: 1087.50 Matches: 216
Percent Similarity: 65.35% Conservative: 99
Best Local Similarity: 44.81% Mismatches: 164
Query Match: 41.59% Indels: 3
DB: 4 Gaps: 3

US-10-669-693-2 (1-504) x US-09-023-655-1056 (1-1854)

QY	9	LeuLeuAlaLeuAlaLeuLeuLeuThrIleAlaLeuSerGlyThrArgAla	28
DB	31	CTTGTGCTCTCTCATCTATTTTCCTTCTCACTCGAGACAGAGCTCT--GGG	87
QY	29	ArgGlyHisLeuProGlyProThrProLeuProLeuGlyAsnLeuLeuGlnLeu	48
DB	88	AGAGAAAACTCCCTCCTGGCCCCACTCCTCCAGTGATTGAATAATCCCTACAGATA	147
QY	49	ArgProGlyAlaLeuTyrsrGlyLeuMetArgLeuSerLysLysTyrsrGlyProValPhe	68
DB	148	GGTATTAAAGACATCAGCAAAATCTTAACAATCTCTCAAAGTCTATGGCCCTGTGTC	207
QY	69	ThrIleTyLeuGlyProTrpArgProValValVallValGlnGlnAlaValArg	88
DB	208	ACTCTGATTATTGGC--CTGAACCACCATAGTGTGCTGCATGATGAAGCAGTAGAAG	264
QY	89	GluAlaLeuGlyGlyGlnAlaGluAlaPheSerGlyArgGlyThrValAlaMetLeuGlu	108
DB	265	GAAGCCCTGATTGATCTTTGGAGAGGAGTTTTCTGGAAGAGGCATTTCCCACCTGGCTGAA	324
QY	109	GlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTrpArgGlnLeuArg	128
DB	325	AGAGCTAACAGGATTGGAATTGTTTCAGCAATGGAAGAATAATGGAAGGAGATCCGG	384
QY	129	LysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeu	148
DB	385	CGTTTCTCCCTCATGCGCTGGGAATTTTGGATGGGGAAGAGAGGAGCATTTGAGAGCCGT	444
QY	149	IleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPhe	168
DB	445	GTTCAAGAGGAAGCCGCTGCTGGAGGATTTGAGAGGATTTGAGAAAACCAAGCCCTCACCTGT	504
QY	169	AspProserLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGly	188
DB	505	GATCCACTTTTCATCTCGGCTGTGCTCCCTGCAATGTGATCTGTCTCCATTTTTCAT	564
QY	189	LeuArgPheSerTyrrGluAspLysGluPheGlnAlaValValArgAlaAlaGlyThr	208
DB	565	AAACGTTTTCATTATAAAGATCAGCAATTTCTTAACCTAATGAAAGAAAGTTGAATGAAAC	624
QY	209	LeuLeuGlyValSerSerGlnGlyGlyGlnThrTyrrGluMetPheSerTrpPheLeuArg	228
DB	625	ATCAAGATTTTGAGAGCCCTGGATCCAGATCTGCAATAATTTTCTCTCATCATGAT	684
QY	229	ProLeuProGlyProHisLysGlnLeuLeuHisValSerThrLeuAlaAlaPheThr	248
DB	685	TACTTCCGGAACTCACACAAATTTACTTAAACAGTTGCTTTTATGAAAAGTTATTATT	744
QY	249	ValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaserGlyProAlaArgAspLeu	268
DB	745	TTGGAAAAAGTAAAGAACACCAAGAAATCAATGGACATCAACACCT--CAGGACTTT	801
QY	269	ValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAlaInsnProGlyThrGluPheThr	288
DB	802	ATTGATTGCTTCTGATGATAAATGGAGAGAAAGAACACCAACCATCTGAAATTTACT	861
QY	289	AsnLyAsnMetLeuMetThrValIleTyrrLeuLeuPheAlaGlyThrMetThrValSer	308
DB	862	ATTGAAAGCTTGGAAAAACATCGCATTGACTTTTGGAGCTGGGACAGACGACCAACG	921
QY	309	ThrThrValGlyTyrrThrLeuLeuLeuMetLysTyrrProHisValGlnLysTrpVal	328
DB	982	CAGGAAGAGATTGAACCTGTGATGGCAGAAACCGGAGCCCTCATCTATTCCTCCCAAG	1041
QY	329	ArgGluGluLeuAsnArgGluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThr	348
DB	982	CAGGAAGAGATTGAACCTGTGATGGCAGAAACCGGAGCCCTCATCTATTCCTCCCAAG	1041
QY	349	ArgLeuProTyrrThrAspAlaValLeuHisGluAlaGlnArgLeuLeuAlaLeuValPro	368
DB	1042	CACATGCCCTACACAGATGCTGTGTGCACGAGGTCACAGATACCTTGACCTTCTCCCC	1101
QY	369	MetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTyrrThrLeuProGln	388
DB	1102	ACCACCTGCCCATGACCTGTGACATTAAATTCAGAAACATATCTCATTCCTCCCAAG	1161
QY	389	GlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePheLysHis	408
DB	1162	GGCAACAACCATTAATTTCCCTGACTTCTGTGTGTACATGACACAAAGAATTTCCCAAC	1221
QY	409	ProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHisGlu	428
DB	1222	CAGAGATGTTTGACCTCATCATCTTCTGGATGAAGGTGGCAATTTTAAGAAAAATGAA	1281
QY	429	AlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGlyLeuAlaLysAla	448
DB	1282	TACTTTCATGCTTCTCAGCAGAAAAACGATTTGTGTGGAGAGCCCTGCGCCGCGATG	1341
QY	449	GluLeuPheLeuPhePheThrThrIleLeuGlnAlaPheSerLeuGluSerProCysPro	468
DB	1342	GAGCTGTTTTTATTCCTGACCTCCATTTTACAGAACTTTAACTGAAATCTCTGGTTGAC	1401
QY	469	ProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAlaPhe	488
DB	1402	CCAAAGAACCTTGACACCACTCCAGTTGTCAATGGTTTTCCTGTGCGGCCCTTCTAC	1461
QY	489	GlnLeu 490	
DB	1462	CAGCTG 1467	

RESULT 11
PCT-US95-05744-4
; Sequence 4, Application PC/TUS9505744
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; APPLICANT: DE MORALS, Sonia M.F.
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
; CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT
; OF S-MEPHNYTOIN METABOLISM
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05744
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,821
; FILING DATE: 06-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/201,118
; FILING DATE: 22-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992


```

; Sequence 15, Application US/08277031B
; Patent No. 6620593
; GENERAL INFORMATION:
; APPLICANT: Hayashi, Koji
; APPLICANT: Sakaki, Toshiyuki
; APPLICANT: Yabusaki, Yoshiyasu
; APPLICANT: Komai, Koichiro
; APPLICANT: Kaneko, Hideo
; APPLICANT: Nakatsuka, Iwao
; TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF
; TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING
; TITLE OF INVENTION: HUMAN CYTOCHROME P450
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5inch, 1.44MB
; COMPUTER: IBM PC
; OPERATING SYSTEM: Dos 5.0
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/277,031B
; FILING DATE: 19-JULY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP-201120/1993
; APPLICATION NUMBER: JP-180246/1993
; APPLICATION NUMBER: JP-208279/1993
; FILING DATE: 20-07-1993
; FILING DATE: 21-07-1993
; FILING DATE: 30-07-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Raymond C. Stewart
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 20-3530P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1473
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-277-031B-15

Alignment Scores:
Pred. No.: 1 5e-108 Length: 1473
Score: 1085.50 Matches: 218
Percent Similarity: 64.94% Conservative: 95
Best Local Similarity: 45.23% Mismatches: 166
Query Match: 41.51% Indels: 3
DB: 4 Gaps: 3

US-10-669-693-2 (1-504) x US-08-277-031B-15 (1-1473)

Qy 9 LeuLeuAlaLeuAlaLeuLeuLeuThrLeuAlaLeuSerGlyThrArgala 28
Db 19 CTGTGTCGTCTCTCATGTTTGCTTCTCTTCACTCTGGAGACAGACTCT--GGG 75
Qy 29 ArgGlyHisLeuProGlyProThrProLeuProLeuGlyAsnLeuLeuInLeu 48
Db 76 AGAGGAAAACTCCCTCCTGGCCCCACTCTCTCCAGTGATTGAAAATATCTTACAGATA 135
Qy 49 ArgProGlyAlaLeuTySerGlyLeuMetArgLeuSerLysTyrGlyProValPhe 68
Db 136 GATATTAGATGTCAGCAAATCTCTTAACCAATCTCTCAAAAATCTATGCCCTGTGTC 195

```


Mon Sep 20 09:09:26 2004

```

|||||... ||| |||||||||||... ||| |||
920 ACAACCTGAGATATGCTCTCCTCTCTGCTGAGACACCCAGAGGTACAGCTAAAGTC 979
Db

329 ArgGluGluLeuAsnArgGluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThr 348
QY :|||||... ||| :||| :|||||... ||| :||| :|||||...
Db CAGGAAGAGATTGAACGTGTGATTGGCAGAACCCGAGCCCTGCTGATCAAGACAGAGC 1039

349 ArgLeuProTyrThrAspAlaValLeuHisGluAlaGlnArgLeuLeuAlaLeuValPro 368
QY :|||||... ||| :||| :|||||... ||| :||| :|||||...
Db CACATGCCCTACACAGATGCTGTGGTCAGCAGAGTCCAGAGATACATTGACCTTCTCCC 1099

369 MetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTyrThrLeuProGln 388
QY :|||||... ||| :||| :|||||... ||| :||| :|||||...
Db ACCAGCCTGCCCATGACGTGACCTGTGACATTAATTCAGAACTATCTCAATCCCAG 1159

389 GlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePheLysHis 408
QY :|||||... ||| :||| :|||||... ||| :||| :|||||...
Db GGCACAACCATATTAATTTCCCTGACTTCTGTGTACATGACACAACAGAAATTTCCCA 1219

409 ProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHisGlu 428
QY :|||||... ||| :||| :|||||... ||| :||| :|||||...
Db CCAGAGATGTTTGACCTTCATCTTCTGGATGAAGTGGCAATTTTAAGAAAGTAAA 1279

429 AlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGlyLeuAlaLysAla 448
QY :|||||... ||| :||| :|||||... ||| :||| :|||||...
Db TACTTCATGCCTTCTCAGCAGAAAACGGATTGTGGGAGAGCCCTGGCCGCGCATG 1339

449 GluLeuPheLeuPhePheThrThrIleLeuGlnAlaPheSerLeuGluSerProCysPro 468
QY :|||||... ||| :||| :|||||... ||| :||| :|||||...
Db GAGCTGTTTTTATCTGACCTCCATTTTACAGAACTTTAACCCTGAAATCTCTGGTTG 1399

469 ProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAlaPhe 488
QY :|||||... ||| :||| :|||||... ||| :||| :|||||...
Db CCAAGAACCTTGACACCACTCCAGTTGTCAATGGATTTCCTCTGTGCGCGCCTTCTAC 1459

489 GlnLeu 490
QY :|||||
Db CAGCTG 1465

```

Search completed: September 18, 2004, 04:35:53
Job time : 153 secs


```
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-669-693-2 (1-504) x US-10-067-668-6 (1-1515)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20
DB 1 ATGGAGGCGACCGGACACCTGGGCGCTGCTGCTGGCGCTGCTGCTGCTGCTGAGC 60
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40
DB 61 CTGGCGCTGCTCGGGACCGAGGCGCGGAGCCACCTGCCCGCGGGCCACCGCGGTACCA 120
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
DB 121 CTGCTGGGAAACCTCTCTGAGCTACGGCCCGGGGCGCTGTATTACGGGCTCATGGGCTG 180
QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValVal 80
DB 181 ACTAAGAAGTACGGACCGGTGTTACCATCTACCTGGGACCGTGGCGGCTGTGTGTGTC 240
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluPheSerGly 100
DB 241 CTGGTTGGGACGAGGCTGTGGGAGGCGCTGGGAGGTCAGGCTGAGGAGTTCAGCGGC 300
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
DB 301 CGGGGAAACCGTAGCATGCTGAGAGGACTTTTGATGGCCATGGGGTTTTCTTCTCCAC 360
QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
DB 361 GGGGAGCGTGGAGGCGAGCTGAGGAAGTTTACCATGCTTGTCTGCGGGACCTGGGCATG 420
QY 141 GlyLysArgGluGlyGlnGluLeuIleGlnAlaGluAlaArgCysLeuValGlnThrPhe 160
DB 421 GGGAGCGAGAAAGCGAGAGCTGATCCAGGGCGAGGCGCGGTGCTGTGGTGAGACATTC 480
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
DB 481 CAGGGGACAGAGGAGCGCCATTCGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAC 540
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
DB 541 GTAGTCTGCTCCCTCTTTGGCCCTCCGCTTCTCTATGAGGATAAGGAGTTCCAGGCG 600
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr 220
DB 601 GTGGTCCGGGCGAGCTGGTGTACCTGTGGGAGTCAGCTCCAGGGGGGTGAGACCTAC 660
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
DB 661 GAGATGTTCTCTGGTTCCTGGGCGCCCTGCCAGGCCCCACAGCAGCTCTCCACCAC 720
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
DB 721 GTCAGCACCTTGGCTGCCCTTACAGTCCGCGAGTGCAGCAGCACCGAGGGAACCTGGAT 780
QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280
DB 781 GCTTGGGGCCCCGACGTCACCTGTGATGTCCTGCTGAAGATGACACAGGAGAA 840
QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
DB 841 CAAACCCAGGCACAGATTCCACCAAGAACATGCTGATGACAGTCATTTATTGCTG 900
QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
DB 901 TTTGCTGGGACGATGAGCGTCAGCACCGCGGTGCTATACCTCTCTCTCTGATGAAA 960
QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
```

RESULT 2

```
US-10-175-696-6
; Sequence 6, Application US/10175696
; Publication No. US20030092658A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Meyers, Rachel
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: NOVEL HUMAN ENZYME FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-193001
; CURRENT APPLICATION NUMBER: US/10/175,696
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 10/067,668
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/266,140
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/823,901
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: PCT/US01/10720
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/193,920
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/862,658
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16380
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/205,675
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/882,837
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19319
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/211,727
; PRIOR FILING DATE: 2000-06-15
```

```

; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-175-696-6

```

Alignment Scores:

Pred. No.:	2.52e-314	Length:	1515
Score:	2615.00%	Matches:	504
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-10-669-693-2 (1-504) x US-10-175-696-6 (1-1515)

[illegible]

Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
Db	841	CAAAACCCAGGACAGAGATTCCACCAAGAACATGCTGATGACAGTCATTTATTGCTG	900
Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuLeuMetLys	320
Db	901	TTTGCTGGGACGATGACGGTCAGCACACCGGTGCGCTATACCTCTCTGCTCTGATGAAA	960
Qy	321	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln	340
Db	961	TACCTCATGTCCAAAGTGGGTACGTGAGGAGCTGATCGGAGCTGGGGCTGGCCAG	1020
Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1021	GCACCAAGCCTAGGGGACCGTACCGCGCTCCCTTACACCGACGCGGTTCGTCATGAGCG	1080
Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
Db	1081	CAGCGGTGCTGGCGCTGTGTGCCCATGGGAATACCCCGACCCCTCATGCGGACCAACCGC	1140
Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1141	TTCCGAGGGTACACCTGCCCCAGGGCAGCGAGGTCTTCCCTCTCTTGGCTCCATCCTG	1200
Qy	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1201	CATGACCCCAACATCTTCAAGACACCCAGAGAGTTCAACCCAGACCGTTTCTGTGATGCA	1260
Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1261	GATGACGGTTCAGAAAGCATGAGCGGTCTCTGCCCTCTCTCTTAGGGAAGCGTGTCTGC	1320
Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrIleLeuGlnAla	460
Db	1321	CTTGGAGAGGCGCTGGCAAAAGCGAGCTCTTCTCTTCTTTCACCAACCATCTCAAGACC	1380
Qy	461	PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1381	TTCTCCCTCGAGAGCCGTCGCCCGCGGACACCCCTGAGCCTCAAGCCACCGTCAGTGGC	1440
Qy	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
Db	1441	CTTTTCAACATTTCCCAAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC	1500
Qy	501	ThrGlnThrArg	504
Db	1501	ACGCAGACCAGA	1512

RESULT 3

US-10-231-913-39

; Sequence 39, Application US/10231913

; Publication NO. US20040005576A1

; GENERAL INFORMATION:

; APPLICANT: Guo, Xiaojia S.

; APPLICANT: Li, Li

; APPLICANT: Patturajan, Meera

; APPLICANT: Shimkets, Richard A.

; APPLICANT: Casman, Stacie J.

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Tchernev, Velizar T.

; APPLICANT: Vernet, Corine A.

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Shenoy, Suresh G.

; APPLICANT: Alsobrook II, John P.

; APPLICANT: Edinger, Schlomit

; APPLICANT: Peyman, John A.

; APPLICANT: Stone, David J.

; APPLICANT: Ellerman, Karen

; APPLICANT: Gangolli, Esha A.

; APPLICANT: Boldog, Ference L.

; APPLICANT: Colman, Steven D.

; APPLICANT: Elsen, Andrew J.

Db 1141 TTCCGAGGGTACACCTGCTGCCAGGCGACGAGGTCTTCCCTCTTGGCTCCATCTG 1200
Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1201 CATGACCCCAACATTTCAAGCACCCAGAAAGATTCAACCCAGACCGTTTCTGGATGCA 1260
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1261 CATGACGCGTTTCAGGAAGCATGAGCGTTTCTCTCCCTTCTTAGGGAGAGGTGTCTGC 1320
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla 460
Db 1321 CTTGAGAGGGGCTGGCAAAAGCGAGGCTTCTCTCTTCCACCAACCATCTACAAAGCC 1380
Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1381 TTTCTCTTGGAGCGCTGGCCGCGGACACCTTGAGCGCTCAAGCCACCGTCACTGAGC 1440
Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1441 CTTTTCACATTTCCCCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1500
Qy 501 ThrGlnThrArg 504
Db 1501 ACGCAGACCAGA 1512

RESULT 8

US-10-120-988-436
; Sequence 436, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt FL_genes Version 2.0
; SEQ ID NO 436
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)...(1523)
US-10-120-988-436

Alignment Scores:

Pred. No.: 4,12e-314 Length: 2085
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-10-669-693-2 (1-504) x US-10-120-988-436 (1-2085)

Qy 1 MetGluAlaThrGlyThrAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuLeuLeuThr 20
Db 9 ATGGAGCGGACCGGACCTGGCGGCTGCTGCTGGCGCTGGCGTGTCTCTCTGCTGCTGACG 68
Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40
Db 69 CTGGCGCTGTCCGGGACCGAGGGCCCGAGGCGACCTGCCCCCGGGCGCCACCGCGGTACCA 128

Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
Db 129 CTGCTGGGAAACCTCTCTGAGCTAGCGCCCGGGCGCTGTATTTCAGGGCTCATCGCGCTG 188
Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValValVal 80
Db 189 AGTAAAGAGTAGCAGGACCGGTGTTCACCATCTACCTGGGACCCCTGGCGGCTGTGGTGGTC 248
Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluLeuPheSerGly 100
Db 249 CTGGTGGGACGAGGCTGTGGGAGGAGCCCTGGGAGGTTCAGGCTGAGGAGTTCAGCGGC 308
Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 309 CGGGGAAACCGTAGCGATCTGGAGGAGCTTTTGATGGCCATGGGGTTCCTCTCTCCAAC 368
Qy 121 GlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuMet 140
Db 369 GGGGAGCGGTGGAGCAGCTGAGGAGTTTACCATGTCTTCTCTGGGACCTGGGCGATG 428
Qy 141 GlyLysArgGlyGluGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db 429 GGGAGCGAGAGGCGAGGAGCTGATCCAGGCGGAGGCCCGGTGTCTGGTGAGACATTC 488
Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
Db 489 CAGGGGACAGAGGAGCGCCATTCGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAC 548
Qy 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db 549 GTAGTCTGCTCTCTCTTTGGGCTCCGCTTCTCTATGAGGATAAGAGTTCCAGGCC 608
Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr 220
Db 609 GTGGTCCGGGCGAGCTGGTGTACCTCTGCTGGAGTACGTCCTCCAGGGGGGTTCAGACCTAC 668
Qy 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
Db 669 GAGATGTCTCTCTGTTCTGCGGCCCTTGCAGGCCCTTGCAGGCCCTCCCAAGCAGCTCTCCACCAC 728
Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db 729 GTACGACCTTGGTGTCTTACAGTCCGGCAGGTGAGCAGCAGCAGGAGGAACTGAT 788
Qy 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
Db 789 GCTTCGGGGCCCGCAGCGTACCTTGTGATGCTTCTGCTGAAGATGGCAGAGGAA 848
Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db 849 CAAACCCAGGACAGAAATTCACCAACAGAACATGCTGATGACAGTCAATTTATTGCTG 908
Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
Db 909 TTTGCTGGGACGATGACGCTCAGCACCGCTCGGCTATACCTCTCTGCTCTCTGATGAAA 968
Qy 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
Db 969 TACCTCATGTCCAAAAGTGGTACGTGAGGAGCTGAATCGGAGGCTGGGGGCTGGCCAG 1028
Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db 1029 GCACCAAGCCTAGGGACCGTACCCGCTCCCTTACACCGACCGGTTCTGCTATGAGCG 1088
Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
Db 1089 CAGCGGCTGCTGGCGTGGTGCCATGGGAATACCCCGCACCTCATGCGGACACCCCGC 1148
Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db 1149 TTCGAGGGGTACACCTGCCCCCAGGCGCAGGAGTCTTCCCTCTCTCTGCTCCATCTCTG 1208
Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420

Db 1209 CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCGATGCA 1268
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1269 GATGACGGTTTCAAGAGCATGAGGGTTCCTGCTTCTCTAGGAGAGCGTGTCTGC 1328
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla 460
Db 1329 CTTGAGAGGGCGCTGGCAAAAGCGAGCTCTTCTCTTCCACCACTCTTCAAGCC 1388
Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1389 TTTCTCCTGGAGAGCGCGTGGCGCGGACCCCTGAGCCTCAAGCCCAACCGTCAGTGC 1448
Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1449 CTTTTCACATTTCCCCAGCCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1508
Qy 501 ThrGlnThrArg 504
Db 1509 ACGCAGACCAGA 1520

RESULT 9

US-10-363-829-102
; Sequence 102, Application US/10363829
; Publication No. US20040142331A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;
; APPLICANT: Altus, Christina M.; Dufour, Gerard E.;
; APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;
; APPLICANT: Jones, Anissa L.; Yu, Jimmy Y.;
; APPLICANT: Wright, Rachel J.; Gietzen, Darryl;
; APPLICANT: Liu, Tommy F.; Yap, Pierre E.;
; APPLICANT: Dahl, Christopher R.; Momiya, Monika G.;
; APPLICANT: Bradley, Diana L.; Rohatgi, Sameer D.;
; APPLICANT: Harris, Bernard; Roseberry Lincoln, Ann M.;
; APPLICANT: Gerstin, Jr., Edward H.; Peralta, Careyana H.;
; APPLICANT: David, Marie H.; Panzer, Scott R.;
; APPLICANT: Flores, Vincent Z.; Daffo, Abel;
; APPLICANT: Marwaha, Rakesh; Chen, Alice J.;
; APPLICANT: Chang, Simon C.; Au, Alan P.;
; APPLICANT: Iman, Rebekah R.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1183 USN
; CURRENT APPLICATION NUMBER: US/10/363,829
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: PCT/US01/27628
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/229,751
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,749
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,750
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,747
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/229,748
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,583
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 60/230,517
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 60/230,610
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 60/230,597
; PRIOR FILING DATE: 2000-09-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PERL Program
; SEQ ID NO 102
; LENGTH: 2604
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: LG:337818.2.2000SEF08
US-10-363-829-102
Alignment Scores:
Pred. No.: 5,82e-314 Length: 2604
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
US-10-669-693-2 (1-504) x US-10-363-829-102 (1-2604)
Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuThr 20
Db 40 ATGAGGCGACCGGACCTGGCGCTGTGCTGGCGCTGGCGCTGCTCTGCTGTGCG 99
Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40
Db 100 CTGGCGCTGTCCGGGACCGAGGCGCCAGGCGCCAGTCCCGCGCGCCAGCGCGCTACCA 159
Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuLeuLeuLeuLeuMetArgLeu 60
Db 160 CTGCTGGGAAACCTCTGCGAGCTACGCGCCCGGGCGCTGTATTACGGGCTCATGCGGCTG 219
Qy 61 SerLysLysTyrglyProValPheThrIleLeuGlyProThrArgProValValVal 80
Db 220 AGTAGAGTAGCGGACCGGTGTACCATCTACCTGGGACCCCTGGCGGCTGTGGTGGTC 279
Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
Db 280 CTGTTGGGAGGAGGCTGTGCGGAGGCGCTGGGAGGCTGGGAGGCTGAGGAGTTCCGGCG 339
Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 340 CGGGAGAACCTAGCATGCTGGAAAGGACTTTTGTATGGCCATGGGCTTTTCTTCTCCAAC 399
Qy 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db 400 GGGGAGCGGTGGAGGAGCTGAGGAAAGTTTACCATCTGCTGCGGGACCTGGGGATG 459
Qy 141 GlyLysArgGluGlyGluGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db 460 GGGAGCGAGAGCGGAGGAGCTGATCCAGCGGAGGCGCGCTGTCTGGTGGAGACATTC 519
Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
Db 520 CAGGGGACAGAGGACGCGCCATTCGATCCCTCTGCTGCGCCAGGCGACCTCCAAC 579
Qy 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrgluAspLysGluPheGlnAla 200
Db 580 GTAGTCTGCTCCCTCTTGGCCCTTCTCTTATGAGGATAAGGAGTTCCAGGCC 639
Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
Db 640 GTGTCGCGGACGCTGGTGTATCCCTGCTGGAGTCACTCCAGGGGGGTTCAGACCTAC 699
Qy 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240
Db 700 GAGATGTTCTCTGTTCTCTGCGGCCCTGCGGCCCTGCGGCCCTGCGGCCCTCTCCACAC 759
Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db 760 GTCAGCACCTGGCTGCTTTCACAGTCCGCGAGGTGAGGAGCACCCAGGGAACCTGGAT 819
Qy 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
Db 820 GCITTCGGGCGCGCACCTGACCTGTGTCGATGCTTCTGCTGAAGATGGCAGGAGGAA 879
Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300

```
Db 880 CAAAACCCAGGCACAGATTCCAAACAAGAACATGCTGATGACAGTATTATTGCTG 939
Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyThrThrLeuLeuLeuMetLys 320
Db 940 TTTCTGGAGCATGACGGTCAGACCAACGGTCGGCTATACCCCTCTCTCTGATGAAA 999
Qy 321 TyrProHisValGlnLysTyrValArgGluGluLeuAsnArgGluLeuGlyValAlaGln 340
Db 1000 TACCTCATGTCCAAAAGTGGTACGTGAGAGCTGAATCCGGAGCTGGGGCTGGCCAG 1059
Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db 1060 GCACCAAGCCTAGGGAGCCGTACCCGCTCTTACACCGACGGCTTCTCATGAGCG 1119
Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrArg 380
Db 1120 CAGGGCTGTGGCGCTGGTCCCATGGAAATACCCGACCCCTCATGCGGACCCCGC 1179
Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db 1180 TTCGAGGGTACACCTGCCCCAGGGCAGGAGTCTTCCCTCTCTTGGCTCCATCCTG 1239
Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1240 CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCGATGCA 1299
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1300 GATGGAGGTTACAGAAAGCATGAGCGTCTCTGCTCTTCTTAGGGAAGCGTCTGC 1359
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
Db 1360 CTGGAGAGGGCTGGCAAAAGCGGAGCTTCTCTCTTCTTCCACCACTCCATCAAGCC 1419
Qy 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1420 TTCTCTCTGAGAGCGCGTCCCGCGCGACACCTGAGCCTCAAGCCACCGTCAGTGGC 1479
Qy 481 LeuPheAsnIleProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1480 CTTTTCAACATTTCCCGAGCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCC 1539
Qy 501 ThrGlnThrArg 504
Db 1540 ACGCAGACCCAGA 1551

RESULT 10
US-10-258-080-16
; Sequence 16, Application US/10258080
; Publication No. US20040029125A1
; GENERAL INFORMATION:
; APPLICANT: Incyte Genomics, Inc.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: BURFORD, Neil
; APPLICANT: RING, Huijun Z.
; APPLICANT: LAL, Preeti G.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: DAS, Debopriya
; APPLICANT: SANJANWALA, Madhusudan M.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: REDDY, Roopa M.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: RAMKIMAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: AU-YOUNG, Janice K.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
```

```
; FILE REFERENCE: PI-0070 USN
; CURRENT APPLICATION NUMBER: US/10/258,080
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/203,509
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/202,234
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/200,185
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/198,403
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US01/11869
; PRIOR FILING DATE: 2001-04-32
; PRIOR APPLICATION NUMBER: US 60/197,590
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 2620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040029125A1 1272843CB1
US-10-258-080-16
```

Alignment Scores:

```
Pred. No.: 5,87e-314 Length: 2620
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
```

US-10-669-693-2 (1-504) x US-10-258-080-16 (1-2620)

```
Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20
Db 41 ATGAGGCGCAGCCGACCTGGCGCTGCTGGCGCTGGCGCTGCTCTGCTGCTGAGC 100
Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProLeuPro 40
Db 101 CTGGCGCTGTCCGGGACAGGCGCCAGGCGCCACCTGCCCCCGGGCCACCGCGCTACCA 160
Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
Db 161 CTGCTGGAAACCTCTCTGAGCTACGCGCGCGGCGCTGTATTTCAGGGCTCATGGCGCTG 220
Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValVal 80
Db 221 AGTAAGAAGTACGACCGCGTGTACCATCTACCTGGGACCTTGGCGGCTGTGGTGGTC 280
Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
Db 281 CTGTTGGCAGAGGCTGTGGGAGGCGCTTGGAGGTCAGGCTGAGGAGTTTCAGCGGC 340
Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 341 CGGGGAACCGTAGCATGCTGGAAGGAGCTTTTGTATGGCCATGGGGTTTCTTCTTCAAC 400
Qy 121 GlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db 401 GGGGAGCGGTGGAGGACGCTGAGGAAGTTTACCATGCTTGTCTCTGGGACCTGGGCGATG 460
Qy 141 GlyLysArgGluGlyGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db 461 GGGAGCGGAGAGGCGAGAGCTGATCCAGGCGGAGGCGCGGTGTCTGTGGAGACATTTC 520
Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180
Db 521 CAGGGGACAGAGGAGCGCCCTTCATCCCTCTCTGCTGCTGGCCAGGCGCCTTCCAAAC 580
Qy 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
```


Db	581	GTAGTCTCTCCCTCTCTTTGGCTTCCTATGAGTAAAGAGTTTCAGGCC	640		APPLICANT: Beresini, Maureen
Qy	201	ValValArgAlaAlaGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr	220		APPLICANT: DeForge, Laura
Db	641	GTGGTCCGGGACGTGGTGGTACCTCTGCTGGAGTCAGTCCACGGGGGTCAGACCTAC	700		APPLICANT: Desnoyers, Luc
Qy	221	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis	240		APPLICANT: Filvaroff, Ellen
Db	701	GAGATGTTCTCTGGTCTCTGGGCGCTGCGAGCCGCCCAAGCAGCTCTCTCCACCAC	760		APPLICANT: Gao, Wei-Qiang
Qy	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuLeuAsp	260		APPLICANT: Gerritsen, Mary E.
Db	761	GTCAGACCTTGGCTTTCAGCTCGGCGAGGTGCAGACACACAGGGAACCTTGAT	820		APPLICANT: Goddard, Audrey
Qy	261	AlaSerGlyProAlaArgAspLeuValAlaPheLeuLeuLysMetAlaGlnGlu	280		APPLICANT: Godowski, Paul J.
Db	821	GCTTCGGGCGCGCAGCTGACCTTGTGATGCTTCTGCTGAAGATGCGACAGGAGAA	880		APPLICANT: Gurney, Austin L.
Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300		APPLICANT: Sherwood, Steven
Db	881	CAAAACCCAGGCACACAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTATTGCTG	940		APPLICANT: Smith, Victoria
Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys	320		APPLICANT: Tumas, Daniel
Db	941	TTTGCTGGACGATGACGCTCAGCACCGTCGGCTATACCTCTCTGCTCTCTGATGAAA	1000		APPLICANT: Watanabe, Colin K
Qy	321	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyGln	340		APPLICANT: Wood, William
Db	1001	TACCTCATGTCCAAAGTGGTACGTGAGGAGTGAATCGGGAGCTGGGGGTGGCCAG	1060		APPLICANT: Zhang, Zemin
Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360		APPLICANT: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
Db	1061	GCACAAAGCTAGGGACCGTACCCGCTCCCTTACACGACGCGTTCCTGATGAGCG	1120		APPLICANT: FILE REFERENCE: P3330R1C345
Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrArg	380		APPLICANT: CURRENT APPLICATION NUMBER: US/10/147,493
Db	1121	CAGCGCTCTGCGCTGGTGGCCATGGGAATACCCCGCACCTCATGCGGACCCCGC	1180		APPLICANT: CURRENT FILING DATE: 2002-05-17
Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400		APPLICANT: Prior Application removed - See File Wrapper or Palm
Db	1181	TTCCGAGGGTACACCTGCCCCAGGCGACGAGGTCTTCCCTCTCTGCTCCATCTCTG	1240		APPLICANT: NUMBER OF SEQ ID NOS: 550
Qy	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420		APPLICANT: SEQ ID NO 63
Db	1241	CATGACCCCAACATCTCAAGCACCCAGAGGATTCACCCAGACCGTTTCTCTGGATGCA	1300		APPLICANT: LENGTH: 2623
Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440		APPLICANT: TYPE: DNA
Db	1301	GATGGACGGTTCAGGAAGCATGAGGGCTTCTGCTTCTTACGGAAGCGTGTCTGC	1360		APPLICANT: ORGANISM: Homo Sapien
Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla	460		APPLICANT: US-10-147-493-63
Db	1361	CTTGGAGAGGCGCTGGCAAAAGCGAGCTCTTCTCTTCTTACCAACCATCTTACAGCC	1420		Alignment Scores:
Qy	461	PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480		Pred. No.: 5,88e-314
Db	1421	TTCTCTCTGAGAGCGCGTCCCGCGGACACCTGAGCCTCAAGCCCAACCGTCAGTGGC	1480		Score: 2615.00
Qy	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500		Percent Similarity: 100.00%
Db	1481	CTTTTCAACATTTCCCGCAGCTTCCAGCTGCAAGTCCGTCCCTTCTGACCTTCTACTCCACC	1540		Best Local Similarity: 100.00%
Qy	501	ThrGlnThrArg	504		Query Match: 100.00%
Db	1541	ACGACAGCAGAGA	1552		DB: 13
RESULT 11				US-10-669-693-2 (1-504) x US-10-147-493-63 (1-2623)	
US-10-147-493-63				Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuThr 20	
; Sequence 63, Application US/10147493				Db 41 ATGGAGGCGACCGGACCTTGGCGCTGTCTGGCGCTGGCGCTGTCTCTGTCTGTACG 100	
; Publication No. US20040029217A1				Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40	
; GENERAL INFORMATION:				Db 101 CTGGCGCTGTCTGGGACACGAGCGCCGAGCCACCTGCCCGGCGCCACGCGCTACCA 160	
; APPLICANT: Baker, Kevin P.				Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60	
				Db 161 CTGCTGGGAAACCTCTGCAGCTACGCGCCCGGGCGCTGTATTTCAGGGCTCATGCGGCTG 220	
				Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValValVal 80	
				Db 221 AGTAAGAAGTACGGACCGGTGTTCACCATCTACCTGGGACCCCTGGCGGCTGTGTGTGTC 280	
				Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100	
				Db 281 CTGTTGGCAGAGGCTGTGGGAGGCGCTTGGAGGTTCAGGCTGAGGAGTTTCAGCGGC 340	
				Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120	
				Db 341 CGGGAAACCTAGCGATGTGGAAAGGACCTTTTGTATGGCCATGGGTTTTTCTTCTCCAC 400	
				Qy 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140	
				Db 401 GGGAGCGGTGGAGGCGAGCTGAGGAAGTTTACCATGCTTGTCTGTGGGACCTGGGATTC 460	
				Qy 141 GlyLysArgGluGlyGluLeuLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160	
				Db 461 GCGAAGCGAAGGCGAGGAGCTGATCCAGCGGAGCGCGCTGTCTGTGTGGAGACATTC 520	
				Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180	

521	Db	CAGGGCACAAGAGACGCCCATTCGATCCCTCCCTGCTGCTGGCCACAGGCCACCTCCAC	581
181	Qy	ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGlnPheGlnAla	200
581	Db	GTAGTCTGCTCCCTCTCTTTGGCTCGGCTTCTCCTATGAGGATNAGGAGTTCAGAGCC	640
201	Qy	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
641	Db	GTGGTCGGGGCAGCTGGTGGTACCTGCTGGAGGTCACTCCCAAGGGGGTCAGACCTAC	700
221	Qy	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis	240
701	Db	GAGATGTTCTCTGGTTCCTGCGGCCCTGCGAGGCCCAACAGACGCTCCTCCACAC	760
241	Qy	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
761	Db	GTACGACACTTGGCTGCTTCACAGTCCGGCAGGTGCAGCAGCACCAGGGGAACCTGGAT	820
261	Qy	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGln	280
821	Db	GCTTCGGGCCCGCAGCGTGACCTTGTGATGCCCTTCCTGCTGGAAGATGGCAACAGAGAA	880
281	Qy	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
881	Db	CAAAACCCAGGCACAGAATTCACCAACAAGAACATGCTGATGACAGCATTTATTGCTG	940
301	Qy	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys	320
941	Db	TTTGCTGGGACGATGACGGTCAGCACCAACCGTTCGGCTATACCTTCCTGCTCCTGATGAA	1000
321	Qy	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln	340
1001	Db	TACCTCATGTCCAAAGTGGTACGTGAGAGCTGAATCGGAGCTGGGGGCTGGCCAG	1060
341	Qy	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
1061	Db	GCACCAAGCCTAGGGACCGTACCCGCTCCCTTACACCGACGGGTTCCTGCATGAGCG	1120
361	Qy	GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrThrArg	380
1121	Db	CAGCGCTCTGGCGCTGGTGGCCCATGGGAATACCCGCACCTCATGCGGACACCCCGC	1180
381	Qy	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
1181	Db	TTCCGAGGTACACCTGCCCCAGAGGCACGAGGTCTTCCCTCCTTGGCTCCATCCTG	1240
401	Qy	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
1241	Db	CATGACCCCAACATCTTCAGACCAACAGAGAGTTCAAACCCAGACCGTTTCCTGGATGCA	1300
421	Qy	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
1301	Db	GATGACCGTTCAGGAGCATGAGGCGTTCCTGCCCTCTTCCCTTGGGAAGCGTGTCTG	1360
441	Qy	LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla	460
1361	Db	CTTGAGAGGGCTGCGAAAAGCGAGCTTCTCTCTTTCACCACTCCTACAGGC	1420
461	Qy	PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
1421	Db	TTCTCCCTGGAGAGCCGGTGGCCGGGACACCTCTAGCCTCAAGGCCACCGTCAGTGGC	1480
481	Qy	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
1481	Db	CTTTTCAACATTCCCCACGCTTCAGCTGGAAGTCCGTCCCACTGACCTTCACTCCACC	1540
501	Qy	ThrGlnThrArg	504
1541	Db	ACGCAGACGAGA	1552

```

; Publication No. US20040033558A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C252
; CURRENT APPLICATION NUMBER: US/10/145,127
; CURRENT FILING DATE: 2002-05-13
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-145-127-63

Alignment Scores:
Pred. No.: 5,98e-314 Length: 2623
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-669-693-2 (1-504) x US-10-145-127-63 (1-2623)

```

Qy	161	GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn	180
Db	521	CAGGGACAGAGAGCGCCATTCGATCCCTCCCTGCTGGCCAGGCCACCTCCAAC	580
Qy	181	ValValCysSerLeuLeuPheGlyLeuAArgPheSerTyrGluAspLysGluPheGlnAla	200
Db	581	GTAGTCTGCTCCCTCTCTTGGCCCTCCGCTTCTCTATGAGGATTAAGGAGTTCCAGGCC	640
Qy	201	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
Db	641	GTGGTCCGGCAGCTGGTGTGTACCTGCTGGAGTTCAGTCCACAGGGGGGTGAGACTTAC	700
Qy	221	GluMetPheSerTrpPheLeuAArgProLeuProGlyProHisLysGlnLeuLeuHisHis	240
Db	701	GAGATGTTCTCTGGTTCCTGCCGCCCTGCCAGGCCCCACAAAGCAGCTCTCTCCACCAC	760
Qy	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
Db	761	GTGACGACCTTGGCTGCTCTTCACAGTCCGCGAGTGCAGCAGCACGAGGGGAACCTGGAT	820
Qy	261	AlaSerGlyProAlaAArgAspLeuValAspAlaPheLeuLeuLysMetalGlnGluGln	280
Db	821	GCITTCGGGCCCGCAGCTGACCTTGTGATGCCCTTCTCTGCTGAAGATGGCACAGGAGAA	880
Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValLleTyrLeuLeu	300
Db	881	CAAAACCCAGGCACAGAATTCCACCAACAAGAACATGCTGATGACAGTCATTTTATTTCGTG	940
Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuLeuMetLys	320
Db	941	TTTCTCTGGGACGATGAGCGGTGAGCACCAACCGTCCGGCTATACCTCTCTGCTTCCTGATGAAA	1000
Qy	321	TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyValAlaGlyGln	340
Db	1001	TACCTCATGTGCCAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGGTGGCCAG	1060
Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1061	GCACCAAGCCTTAGGGGACCGTACCCGCGCTCCCTTTACACCGACGCGGTTCTGCGATGAGCGG	1120
Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
Db	1121	CAGCGCTGCTGGCGCTGGTGCCCATGGGAATACCCCGCACCTCATGCGGACCAACCCGC	1180
Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1181	TTCCGAGGGTACACCTGCCCCAGGCACGAGGAGTCTTCCCTCTCTTGGTCTCCATCCTG	1240
Qy	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1241	CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCCGTTTCTCTGATGCA	1300
Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1301	GATGACCGGTTTCAGGAACATGAGGCGTCTCTGCCCTTCTCTTAGGAAGCGGTGTCTGC	1360
Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAla	460
Db	1361	CTTGGAGAGGGCTGGCAAAAGCGGAGCTCTTCTCTTCTTCAACCAACATCTCAAGGCC	1420
Qy	461	PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1421	TTCTCCCTTGGAGAGCCCGTGCCGCCGAGACACCTTGAGCTTCAGCCCAACCGTCAGTGGC	1480
Qy	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
Db	1481	CTTTTCAACATTTCCCCAGCCTTCCAGTGCAGTCCGTAAGTCCGTCCCACTGACTTCACTCCACC	1540
Qy	501	ThrGlnThrArg	504
Db	1541	ACGCGACAGA	1552

```

RESULT 13
US-10-160-503-63
; Sequence 63, Application US/10160503
; Publication No. US20040033559A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
;
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; ACIDS ENCODING THE SAME
;
; FILE REFERENCE: P3330R1C446
; CURRENT APPLICATION NUMBER: US/10/160,503
; CURRENT FILING DATE: 2002-05-30
;
; NUMBER OF SEQ ID NOS: 550
;
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
;
; US-10-160-503-63

```

Alignment Scores:		
Pred. No.:	5.89e-314	2623
Score:	2615.00	Matches: 504
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
Dg.	13	Gaps: 0

Db 461 GGGAGCGAGAGAGCGAGAGCTGATCCAGCGGAGCGCCGGGTCTGTGGTGGAGACATTC 520
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
Db 521 CAGGGGACAGAGGAGCGCCATTCGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAAC 580
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db 581 GTAGTCTGCTCCTCTCTTTGGCCTCCGCTTCTCTATGAGGATAAGAGATTCCAGGCC 640
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr 220
Db 641 GTGGTCCGGGAGCTGGTGGTACCTGCTGGAGTCAAGTCCAGGGGGTTCAGACCTAC 700
QY 221 GluMetPheSerTyrPheLeuArgProGlyProGlyProHisLysGlnLeuLeuHis 240
Db 701 GAGATGTTCTCTGCTGTTCTGCGGCCCCCTGCGAGGCCCCCAACAGCAGCTCTCCACCAC 760
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db 761 GTCAGCACCTTGGTGCCTTCACAGTCCGCGAGGTGCAGCAGCACACAGGGGAACCTGGAT 820
QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
Db 821 GCTTCGGGCCCCGACGCTGATGCTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880
QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db 881 CAANAACCCAGCAGAGATTCCACAACAGAACATGCTGATGACAGTCAATTTATTTGCTG 940
QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
Db 941 TTTGCTGGGAGATGACGGTCAGCACACGCTCGGCTATACCTTCTGCTCCTGCTGCTGCTGCT 1000
QY 321 TyrProHisValGlnLysTyrValArgGluLeuLeuAsnArgGluLeuGlyValAlaGlyGln 340
Db 1001 TACCCTCATGTCCAAAAGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060
QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db 1061 GCACCAAGCTAGGGAGCGGTACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1120
QY 361 GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrThrArg 380
Db 1121 CAGCGGCTGCTGGCTGGTGGCCATCGGAATACCCCGCACCCTCATGCGGACCCCGC 1180
QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db 1181 TTCGAGGGTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1240
QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1241 CATGACCCCAACATCTTCAAGACCCAGAGAGTTCACCCAGACCGTTTCTCTGGATGCA 1300
QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1301 GATGGACGGTTCAGGAAGCATGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1360
QY 441 LeuGlyGlnGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla 460
Db 1361 CTTGGAGAGGGCTGGCAAAAGGGAGCTCTTCTCTTCTTCAACCACTCTTACAGGCC 1420
QY 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1421 TTTCTCCTGGAGGCCGCTGCGCGCGGACACCTGAGCTTCAAGCCCAACCGTCACTGGC 1480
QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1481 CTTTTCACATTTCCCGCAGCTTCCAGCTGCAAGTCCGCTCCCACTGACCTTCACTCCACC 1540
QY 501 ThrGlnThrArg 504

Db 1541 ACGCAGACCAGA 1552
RESULT 14
US-10-143-118-63
; Sequence 63, Application US/10143118
; Publication No. US2004003835A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/143,118
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-143-118-63
Alignment Scores:
Pred. No.: 5,88e-314 Length: 2623
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-10-669-693-2 (1-504) x US-10-143-118-63 (1-2623)
QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuThr 20
Db 41 ATGGAGCGGACCGGACCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 100
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProThrProLeuPro 40
Db 101 CTGGCGCTGTCCGGGACCGAGGCCCGGAGCCCTGCCCCCGGCGCCACCGCGCTACCA 160
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyValAlaLeuTyrSerGlyLeuMetArgLeu 60
Db 161 CTGCTGGGAAACCTCTCGAGCTACGCGCCCGGGCGCTGTATTCAGGGCTCATGCGGCTG 220
QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValValVal 80
Db 221 AGTAAGAAGTACGAGACCGGTTCACCATCTACCTGGGACCTGGCGCTGCTGCTGCTGCTGCT 280
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGlnGluPheSerGly 100
Db 281 CTGGTTGGCAGAGGCTGTGGGAGGCCCTGGGAGGTTCAGGCTGAGGAGTTTCAGCGGC 340
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 341 CGGGAAACCGTAGCATGCTGGAAGGGACTTTTGTATGACCATGGGGTTTTTCTTCTTCTTCTTCT 400
QY 121 GlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140

Db 401 GGGAGAGCGGTGGAGGACGCTGAGGAGTTTACCATGCTTGTCTCTGCGGAGCCTGGGCATG 460
Qy 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db 461 GGGAGCGAGAGCGGAGGAGCTGATCAGCGGAGGCGCGGTGTCTGGTGGAGACATTC 520
Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
Db 521 CAGGGACAGAGGAGCGCCATTCGATCCCTCCCTGCTGCTGGCCAGCGCCACCTCCAAAC 580
Qy 191 ValValCysSerLeuLeuPheCysLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db 581 GTAGTCTCTCCCTCTTTGGCTCTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 640
Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
Db 641 GTGGCTGGCGGAGCTGGTGGTACCTGCTGGAGTCAGCTCCAGGGGGGTTCAGACCTAC 700
Qy 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
Db 701 CAGATGTTCTCTCTGTTCTCTGGGCGCTCCAGGCGCCCAAGCAGGCTCTCTCCACCAC 760
Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db 761 GTCAGACCTTGGCTGGCTTTCAGCTCCGAGGTGCAGCAGCACCAGGGGAACCTGGAT 820
Qy 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGln 280
Db 821 GCTTCGGGCGCGCAGCTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880
Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db 881 CAAAACCCAGGACACAGAAATTCACCAACAAAGAACATGCTGATGACGTCAATTTATTGCTG 940
Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
Db 941 TTTGCTGGAGCATGAGGTGACGACACACGCTGGCTATACCTCTCTCTCTCTCTCTCTCT 1000
Qy 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGlnLeuGlyAlaGlyGln 340
Db 1001 TACCCTCATGTCCAAAGTGGTACGTGAGGAGTGAATCGGGAGCTGGGGGTGGCCAG 1060
Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db 1061 GCACCAAGCCTAGGGACCGGTACCCGCTCCCTTACACCGACCGGTTCGTGATGAGCG 1120
Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
Db 1121 CAGCGCTGCTGGGCTGGTGGCTATGGGAATACCCCGACCTCATGCGGACCAACCCGC 1180
Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db 1181 TTCGAGGGTACACCTCGCCCGGACGCGAGGTCTTCCCTCTCTCTCTCTCTCTCTCTCT 1240
Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1241 CATGACCCCAACATCTTCAAGACACCCAGAGAGTTCAACCCAGACCGTTTCTTGGATGCA 1300
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1301 GATGACGGTTCAGGAAGCATGAGCGTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1360
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAla 460
Db 1361 CTTGGAGAGGCGCTGGGAAAAGCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1420
Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1421 TTTCTCTGGAGAGCGCGCTGGCGGACACCTTGAGCCTCAAGCCACCGCTCAGTGGC 1480
Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1481 CTTTTCACATTTCCCCCAGCCTTCCAGCTGCAAGTCCGTCCTGCCACTGACCTTCACTCCACC 1540

Qy 501 ThrGlnThrArg 504
Db 1541 ACGCAGACCAGA 1552
RESULT 15
US-10-144-993-63
; Sequence 63, Application US/10144993
; Publication No. US20040038336A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C261
; CURRENT APPLICATION NUMBER: US/10/144,993
; CURRENT FILING DATE: 2002-05-13
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 63
; LENGTH: 2623
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-144-993-63
Alignment Scores: Length: 5,88e-314 2623
Pred. No.: 2615.00 504
Score: 2615.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 13 Gaps: 0
US-10-669-693-2 (1-504) x US-10-144-993-63 (1-2623)
Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr 20
Db 41 ATGAGGCGGACCGGACCTGGGCGCTGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 100
Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProThrProLeuPro 40
Db 101 CTGGCGCTGTCTCGGACCCAGGCGCCAGGCGCACCTGCCCCCGGCGCCACCGCGCTACCA 160
Qy 41 LeuLeuGlyAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
Db 161 CTGCTGGAAACCTCTCTGCGAGTACGCGCCCGGCGCGCTGATTTCAGGGCTCATCGCGGCTG 220
Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValValVal 80
Db 221 AGTAAGAAGTACGAGCCCGGTTTCAACCATCTACTGGGACCTTGGCGCGCTGCTGCTGCTG 280
Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlnAlaGluGluPheSerGly 100
Db 281 CTGCTGGGAGGAGGCTGTGCGGAGGCGCTTGGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 340
Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 341 CGGGGAACCGTAGCATGCTGGGAAGGACCTTTGATGGCCATGGGGTTTCTTCTCTCCAAC 400

```
QY 121 GlyGluArgTTPArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db |||||
QY 401 GGGGAGCGGTGGAGCAGCTGAGGAGATTACATGCTTGCTCTCGGGACCTGGGCATG 460
Db |||||
QY 141 GlyLysArgGluGluGluLeuLeGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db |||||
QY 461 GGGAAAGCGAAGGCGAGGAGCTGATCCAGGCGGAGGCCCGGTGCTGGTGAGACATTC 520
Db |||||
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
Db |||||
QY 521 CAGGGGACAGAGGACGCCCATTCGATCCCTCCCTGCTGTCGCCACAGCCACCTCCAAAC 580
Db |||||
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db |||||
QY 581 GTAGTCTGCTCCCTCTCTTTGGCTTCGCTCTCTATGAGGATAAGGAGTTCAGGCC 640
Db |||||
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
Db |||||
QY 641 GTGGTCCGGGCGAGCTGGTGGTACCTGCTGGGAGTCAGCTCCAGGGGGGTCAGACCTAC 700
Db |||||
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240
Db |||||
QY 701 GAGATGTTCTCTCTGTTCTCGGCCCTGCGAGCCCTGCGAGCCGCCCAAGCAGCTCTCCACCAC 760
Db |||||
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db |||||
QY 761 GTCAGCACCTTGGCTGCTTCACAGTCCGGCAGGTGTCAGCAGCACCCAGGGGAACCTGGAT 820
Db |||||
QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280
Db |||||
QY 821 GCCTTCGGGCGCCCGCACCTGACCTGTGATGCCCTTCCTGTGAAGATGGCACAGAGGAA 880
Db |||||
QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db |||||
QY 881 CAAACCCAGCAGCAGATTCACCAACAAGAACATGCTGATGACAGTCATTTATTGCTG 940
Db |||||
QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
Db |||||
QY 941 TTTGCTGGGACGATGACGGTCAGCACCCAGCTCGGCTATACCCCTCTGCTCTGATGAAA 1000
Db |||||
QY 321 TyrProHisValGlnLysTrpValArgGluLeuAsnArgGluLeuGlyValaGlyGln 340
Db |||||
QY 1001 TACCCCTCATGTCAAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGGCTGGCCAG 1060
Db |||||
QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db |||||
QY 1061 GCACCAAGCCCTAGGGACCGTACCCGCTCCCTTACACCGACGCGTTCCTGCATGAGCG 1120
Db |||||
QY 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
Db |||||
QY 1121 CACCGGCTGCTGGCGCTGGTGCCATGGGAAATACCCCGCACCCCTCATCGCGACCCCGC 1180
Db |||||
QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db |||||
QY 1181 TTCGAGGTTACACCTCGCCCGCCAGGCGACGAGTCTTCCCTCTCTTGCCTCATCTCTG 1240
Db |||||
QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db |||||
QY 1241 CATGACCCCAACATCTCAAGACCCAGAGAGTTCAACCCAGACCGTTTCTCTGGATGCA 1300
Db |||||
QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db |||||
QY 1301 GATGGACGGTTACAGAGCATGAGGGGTTCTGCCCCCTTCTCTTAGGGAAGCGGTCTGC 1360
Db |||||
QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
Db |||||
QY 1361 CTTGGAGAGGGCTGGCAAAAGCGAGCTCTTCTCTTCTTACACCATCTCTACAGGCC 1420
Db |||||
QY 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db |||||
QY 1421 TTCTCTCTGAGAGCCCGTGGCCCGGACACCTGAGCCCTCAAGCCCAACCGTCACTGGCC 1480
Db |||||
```

```
QY 481 LeuPheAsnIleProProAlaPheGlnLeuValArgProThrAspLeuHisSerThr 500
Db |||||
QY 1481 CTTTTCAACATTTCCCCAGCCTTCCAGCTGCAAGTCCGTCCCACTGACCTTCACTCCACC 1540
Db |||||
QY 501 ThrGlnThrArg 504
Db 1541 ACGCAGACCAGA 1552
```

Search completed: September 18, 2004, 06:17:48
Job time : 777 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 17, 2004, 18:58:32; Search time 635 Seconds
(without alignments)
3371.798 Million cell updates/sec

Title: US-10-669-693-2
Perfect score: 2615
Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10669693/runat.15092004.101309.5839/app.query.fasta.1.647
-DB=N Geneseq 29Jan04 -QFMR=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10669693 @CGN 1.1.470 @runat.15092004.101309.5839 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N Geneseq 29Jan04:*

1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002s:.*
7: geneseqn2003as:.*
8: geneseqn2003bs:.*
9: geneseqn2003cs:.*
10: geneseqn2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2615	100.0	1515	6	ABT05472 DNA of NO
2	2615	100.0	1927	7	ACA61907 cDNA enco
3	2615	100.0	1927	9	AAD60552 Human cyt
4	2615	100.0	2085	7	ABX71208 Novel hum
5	2615	100.0	2604	6	ABQ72550 Human MDD
6	2615	100.0	2620	6	AAD24671 Human dru
7	2615	100.0	2623	4	AAS21275 Human cDN
8	2615	100.0	2623	7	ACD23884 Novel hum

9	2615	100.0	2623	7	ACA67025	ACA67025 cDNA enco
10	2615	100.0	2623	7	ACA03634	ACA03634 cDNA enco
11	2615	100.0	2623	7	ABX89172	ABX89172 DNA enco
12	2615	100.0	2623	7	ACD41826	ACD41826 Human sec
13	2615	100.0	2623	7	ACA04055	ACA04055 Human cDN
14	2615	100.0	2623	8	ADA45582	ADA45582 Novel hum
15	2615	100.0	2623	8	ADA76013	ADA76013 Human PRO
16	2615	100.0	2623	8	ADA18663	ADA18663 Human PRO
17	2615	100.0	2623	8	ADA61286	ADA61286 Homo sapi
18	2615	100.0	2623	8	ADB19071	ADB19071 Novel hum
19	2615	100.0	2623	8	ADB27612	ADB27612 cDNA enco
20	2615	100.0	2623	8	ADA86091	ADA86091 Novel hum
21	2615	100.0	2623	8	ADB15655	ADB15655 Human PRO
22	2615	100.0	2623	8	ADA47441	ADA47441 Human PRO
23	2615	100.0	2623	8	ADA67236	ADA67236 Human PRO
24	2615	100.0	2623	8	ADB30243	ADB30243 cDNA enco
25	2615	100.0	2623	8	ADA85539	ADA85539 Novel hum
26	2615	100.0	2623	8	ADA86751	ADA86751 Human PRO
27	2615	100.0	2623	8	ADA79055	ADA79055 Human PRO
28	2615	100.0	2623	8	ADA87194	ADA87194 Novel hum
29	2615	100.0	2623	8	ADB16396	ADB16396 Human PRO
30	2615	100.0	2623	8	ADA91488	ADA91488 Novel hum
31	2615	100.0	2623	8	ADB14551	ADB14551 Human PRO
32	2615	100.0	2623	8	ADB18512	ADB18512 Novel hum
33	2615	100.0	2623	8	ADA93727	ADA93727 Human PRO
34	2615	100.0	2623	8	ADB19623	ADB19623 Novel hum
35	2615	100.0	2623	8	ACD98455	ACD98455 Novel hum
36	2615	100.0	2623	8	ADA74189	ADA74189 Human PRO
37	2615	100.0	2623	8	ADA74189	ADA74189 Human PRO
38	2615	100.0	2623	8	ADB24422	ADB24422 Human PRO
39	2615	100.0	2623	8	ADA81946	ADA81946 Human PRO
40	2615	100.0	2623	8	ADA74909	ADA74909 Human PRO
41	2615	100.0	2623	8	ADA84987	ADA84987 Novel hum
42	2615	100.0	2623	8	ADA84435	ADA84435 Novel hum
43	2615	100.0	2623	8	ADB29691	ADB29691 cDNA enco
44	2615	100.0	2623	8	ADA80219	ADA80219 Human PRO
45	2615	100.0	2623	8	ADA75461	ADA75461 Human PRO

ALIGNMENTS

RESULT 1

ABT05472

ID ABT05472 standard; DNA; 1515 BP.

XX

AC ABT05472;

XX

DT 11-OCT-2002 (first entry)

XX

DE DNA of NOVX 16 SEQ ID No 39.

XX

Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipaeamic;
neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
tranquiliser; neuroleptic; antidiabetic; antitumor; antiinflammatory;
anti-HIV; antiallergic; antirheumatic; antithrombotic; NOX; diabetes;
metabolic disorder; obesity; infectious disease; Alzheimer's disease;
anorexia; neurodegenerative disorder; Parkinson's disease; obesity;
immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
metabolic syndrome X; wasting disorder; cancer; neurological disorder;
epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
vesicular transport; cystic fibrosis; gastrointestinal disorder;
diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
multiple sclerosis; rheumatoid arthritis; transgenic animal;
gene therapy; gene; ds.

XX Unidentified.

OS

XX WO200246409-A2.

PN

XX 13-JUN-2002.

PD

PF 06-DEC-2001; 2001WO-US046586.

XX

PR 06-DEC-2000; 2000US-0251660P.
 PR 12-DEC-2000; 2000US-0255029P.
 PR 08-JAN-2001; 2001US-0260326P.
 PR 24-JAN-2001; 2001US-0263800P.
 PR 20-FEB-2001; 2001US-0265942P.
 PR 24-APR-2001; 2001US-0286183P.
 PR 20-AUG-2001; 2001US-0313627P.
 PR 12-SEP-2001; 2001US-0318712P.
 XX (CURA-) CURAGEN CORP.
 FA
 XX Guo X, Li L, Patturajan M, Shimkets RA, Casman SJ, Malyankar UM;
 PI Tchernev VT, Vernet CAM, Spytek KA, Shenoy SG, Alsobrook JP;
 PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA, Boldog FL;
 PI Colman SD, Eisen AJ, Liu X, Padigar M, Spaderna SK, Zerhusen BD;
 XX
 DR WPI; 2002-547774/58.
 DR P-PSDB; ABJ04659.
 DR
 XX Novel isolated polypeptide, designated NOVX, useful for treating or
 PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
 PT metabolic, neurodegenerative, immune and hematopoietic disorders.
 XX
 PS Claim 9; Page 148; 42pp; English.
 CC
 CC The invention relates to an isolated polypeptide, designated NOVX,
 CC comprising a sequence fully defined in the specification. The isolated
 CC protein, its encoding polynucleotide or an antibody created from the
 CC protein is useful in the manufacture of a medicament for treating a
 CC syndrome associated with a human disease, preferably a NOVX-associated
 CC disorder, or for treating or preventing a NOVX-associated disorder in a
 CC subject, preferably human. The isolated protein, its encoding
 CC polynucleotide or an antibody created from the protein are also useful
 CC for treating or preventing metabolic disorders, diabetes, obesity,
 CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
 CC disease, Parkinson's disorder, immune disorders, hematopoietic
 CC disorders, and various dyslipidemias, metabolic disturbances associated
 CC with obesity, the metabolic syndrome X, wasting disorders associated with
 CC chronic diseases, and cancer. The isolated protein, its encoding
 CC polynucleotide or an antibody created from the protein are useful for
 CC treating or preventing neurological disorders such as epilepsy, stroke,
 CC mental disorders including mood, anxiety, schizophrenic disorders,
 CC disorders of vesicular transport such as cystic fibrosis, diabetes
 CC mellitus, goiter, gastrointestinal disorders including ulcerative
 CC colitis, other conditions associated with abnormal vesicle trafficking
 CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
 CC arthritis. A cell comprising the vector of the invention is useful for
 CC producing non-human transgenic animals. The polynucleotide of the
 CC invention can be used to treat disorders by gene therapy. This
 CC polynucleotide sequence represents the DNA encoding one of the isolated
 CC NOVX proteins of the invention
 XX
 SQ Sequence 1515 BP; 274 A; 492 C; 460 G; 289 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,546-231 Length: 1515
 Score: 2615.00 Matches: 504
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps:

US-10-669-693-2 (1-504) x ABT05472 (1-1515)

Db 121 CTGCTGGAAACCTCTCTGAGCTACGGCCCGGGGGCTGTATTTCAGGGCTCATCGGCTG 180
 Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTprArgProValValVal 80
 Db 181 AGTAAGAAGTACGACCGGTGTTCCACCATCTACCTGGGACCGTGGCGGCTGTGGTGTGTC 240
 Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
 Db 241 CTGGTTGGGACGAGGCTGTGCGGAGGCCCTGGAGGTTCAGGCTGAGAGTTCAGCGGC 300
 Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
 Db 301 CGGGGAACCGTAGCATGCTGAAGAGGACTTTTATGGCCATGGGGTTTCTCTCCCAAC 360
 Qy 121 GlyGluArgTprArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
 Db 361 GGGGAGCGGTGGAGGACGCTGAGGAAGTTTACCATGCTTGTCTGCGGGACTGGGCAATG 420
 Qy 141 GlyLysArgGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
 Db 421 GGGAAAGCCAGAGGAGGAGAGCTGATCCAGGGGAGGCGCCGCTGTCTGTGGAGACATTC 480
 Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
 Db 481 CAGGGGACAGAGGAGCGCCATTCGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAC 540
 Qy 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
 Db 541 GTAGTCTCTCCCTCTCTTTGGCTCCGCTTCTCTATGAGGATAGAGGATTCACAGGCC 600
 Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
 Db 601 GTGGTCCGGGAGCTGGTGGTATCCCTGCTGGGAGTCAGCTCCAGGGGGGTTCAGACTAC 660
 Qy 221 GluMetPheSerTprPheLeuArgProLeuArgProGlyProGlyProHisLysGlnLeuHis 240
 Db 661 GAGATGTTCTCTGTTCTGCGGCCCTGCCAGGCCGCCCAAGCAGCTCTCTCCACCAC 720
 Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
 Db 721 GTGAGCACCTTGGCTGCTTCACGTCGCGGAGGTGACAGCACCCAGGGGAACTGGAT 780
 Qy 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280
 Db 781 GCTTGGGGCCCGCAGCTGACCTGTGTCGATGCTCTCTGAAAGATGGCACAGGAGGAA 840
 Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
 Db 841 CAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTATTTGCTG 900
 Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
 Db 901 TTTGCTGGGACGATGACGCTCAGCACCGGTGGGTATACCTCTCTCTCTGATGATGAAA 960
 Qy 321 TyrProHisValGlnLysTprValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
 Db 961 TACCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGTGGGCAG 1020
 Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
 Db 1021 GCACCAAGCTAGGGGACCGTACCCGCTCTCTTACACCGACGGGTTCGATGAGGCG 1080
 Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
 Db 1081 CAGCGGTGCTGGCGTGGTGGCCATGGGAATACCCCGCCACCTCATGCGGACCCACCGCG 1140
 Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
 Db 1141 TTCGAGGGTACACCTGCGCCCGGACGAGGAGTCTTCCCTCTCTCTGCTGCTCATCTG 1200
 Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420

Db 1201 CATGACCCCAACATCTTCAAGCACCAGAGAGTTCAACCCAGACCGTTTCTCGATGCA 1260
 QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyValCys 440
 Db 1261 CATGACCGTTTCAGGAACATGAGCGTTCCTGCCCCCTTCTCTTAGGGAAGCGTGTCTGC 1320
 QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAla 460
 Db 1321 CTTTGAGAGCGCTGGCAAAAGCGAGCTCTCTCTTCTTCCACCACTCTACAGCC 1380
 QY 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
 Db 1381 TTTCTCCCTGGAGAGCCGTCGCCCGGACACCTCTGAGCCTCAAGCCACCGTCAGTGGC 1440
 QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
 Db 1441 CTTTTCACATTCCCCAGCCTTCCAGCTGCAGTGCAGTCCGCTCCCACTGACCTTCACTCCACC 1500
 QY 501 ThrGlnThrArg 504
 Db 1501 ACGCAGACCAGA 1512
 RESULT 2
 ID ACA61907
 XX ACA61907 standard; cDNA; 1927 BP.
 AC ACA61907;
 XX
 XX 28-JUL-2003 (first entry)
 XX cDNA encoding human cytochrome P450, 33303.
 XX Human; ss; gene; cytochrome; P450; cellular proliferation; cancer;
 KW cellular differentiation; cell adhesion; cell motility; cell migration;
 KW inflammation; autoimmune disorder; rheumatoid arthritis; diabetes;
 KW multiple sclerosis; systemic lupus erythematosus; conjunctivitis; asthma;
 KW allergy; Grave's disease; graft-versus-host disease; vascular disorder;
 KW central nervous system disorder; Alzheimer's disease; anxiety disorder;
 KW amnesia; depression; migraine; Parkinson's disease; obesity; cachexia;
 KW cell-signaling-associated disorder; metabolism-associated disorder;
 KW anorexia; steroid-associated disorder; fatty acid-associated disorder.
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1. .1515
 FT CDS /*tag= a
 FT /*product= "33303"
 FT /note= "Cytochrome P450"
 FT sig_peptide 1. .87
 FT /*tag= b
 FT /note= "Signal sequence"
 FT mat_peptide 88. .1512
 FT /*tag= c
 FT /note= "Mature 33303"
 XX
 XX US2003022334-A1.
 PN
 XX
 XX 30-JAN-2003.
 PD
 XX
 XX 04-FEB-2002; 2002US-00067668.
 PF
 XX
 XX 02-FEB-2001; 2001US-0266140P.
 PR
 XX (GLUC/) GLUCKSMANN M A.
 PA
 XX
 XX Glucksmann MA;
 PI
 XX
 XX WPI; 2003-447427/42.
 DR P-PSDB; ABU10222.
 XX
 XX New 33312, 33303 or 32579 nucleic acid molecule encoding cytochrome P450
 PT polypeptides, useful for diagnosing and treating cancer, inflammation,

PT vascular disorders, CNS disorders or metabolic disorders, and in
 FT pharmacogenomics.
 XX
 XX Claim 1; Page 47; 76pp; English.
 CC The invention relates to an isolated nucleic acid molecule comprising any
 CC of the 6 nucleic acid sequences, given in the specification which encodes
 CC a polypeptide having any of the 3 amino acid sequences, given in the
 CC specification. The nucleic acid molecule, protein and antibody are useful
 CC in screening assays and predictive medicine, such as in diagnostic
 CC assays, prognostic assays, monitoring clinical trials and
 CC pharmacogenetics. The composition and method are used in diagnosing and
 CC treating 33312, 33303 or 32579 cytochrome P450-related disorders, such as
 CC cellular proliferative and/or differentiative disorders (e.g. cancer),
 CC disorders associated with undruggable or deficient cell adhesion,
 CC motility or migration, inflammatory or autoimmune disorders (e.g.
 CC rheumatoid arthritis, diabetes, multiple sclerosis, systemic lupus
 CC erythematosus, conjunctivitis, allergies, asthma, Grave's disease or
 CC graft-versus-host disease), vascular disorders, central nervous system
 CC (CNS) disorders (e.g. Alzheimer's disease, anxiety disorders, amnesia,
 CC depression, migraine or Parkinson's disease), cell-signaling-associated
 CC disorders, metabolism-associated disorders (e.g. obesity, cachexia or
 CC anorexia), steroid-associated disorders and fatty acid-associated
 CC disorders. The present sequence represents cDNA encoding the human
 CC cytochrome P450, 33303
 XX
 SQ Sequence 1927 BP; 383 A; 625 C; 550 G; 369 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3 53e-231 Length: 1927
 Score: 2615.00 Matches: 504
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0

US-10-669-693-2 (1-504) x ACA61907 (1-1927)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr 20
 Db 1 ATGGAGCGCAGCCGACCTGGCGCTGCTGCTGGCGCTGGCGCTGCTGCTGCTGCTGACG 60
 QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40
 Db 61 CTGGCGCTGTCCGGGACCCAGGCGCCAGGCGCACCTGCCCGCGCGCCAGCGCTACCA 120
 QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTySerGlyLeuMetArgLeu 60
 Db 121 CTGCTGGGAAACCTCTCTGAGCTACGCGCGCGCGCGCTGTATTCAGGCGCTCATGCGGCTG 180
 QY 61 SerLysLysTyThrGlyProValPheThrIleTyLeuGlyProThrArgProValValVal 80
 Db 181 AGTAAGAAGTACGACCGCGTTCACCATCTACCTGGGACCGTGGCGCGCTGTGGTGGTC 240
 QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluPheSerGly 100
 Db 241 CTGGTTGGCGAGGCGTGTGGGAGGCGCTGGGAGGCGCTGAGGAGTTCAGCGGCTCAGCGGC 300
 QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPheSerAsn 120
 Db 301 CGGGAAACCGTAGCGATCTGAGAGGAGCTTTTGATGGCCATGGGGTTTCTTCTCCAAC 360
 QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
 Db 361 GGGGAGCGGTGGAGCGACCTGAGGAGTTTACCATGCTTCTCTCGGGACCTCGGGCATG 420
 QY 141 GlyLysArgGluGlyGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
 Db 421 GGAAGCGAGAAGCGAGGAGCTGATCCAGCGGAGGCGCGCGTGTCTGTGTGAGACATTC 480
 QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
 Db 481 CAGGGGACAGAAGGACGCCCATTCCTCTCTGCTGCTGGCGCGCGCCACCTCCAAC 540

Mon Sep 20 09:09:26 2004

us-10-669-693-2.rng

181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
 541 GTAGTCTGCTCCCTCTTTGGCTCCGCTTCTCTATGAGGATAAGGAGTTCCAGGCC 600
 201 ValValArgAlaAlaGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
 601 GTGGTCCGGGCGAGTGGTGTGACCTGCTGGAGTCACTCCAGGGGGGFCAGACCTAC 660
 221 GluMetPheSerTrpLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240
 661 GAGATGTCTCTCTGCTTCTCGGGCCCTGACAGGCCCCCAAGACGAGCTCTCCACCAC 720
 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
 721 GTGAGCACCTTGGCTGGCTTCCAGTCCGCGAGGTGCAGCAGCACAGGGGAACTGGAT 780
 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280
 781 GCTTCGGGCCCGCAGCGTACCTTGTGATGCTTCTGCTGAAGATGGCAGAGGAGAA 840
 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
 841 CAAGAACCAGGACAGCAANTTCCCAACAGAACATGCTGATGACAGTCAATTTATTGCTG 900
 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
 901 TTTGCTGGGACGATGACGTCAGCAGCACCGTGGCTATACCTCTCTCTCTGATGAAA 960
 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyValAlaGlyGln 340
 961 TACCCTCATGTCCAAAGTGGGTAGCTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAG 1020
 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
 1021 GACCAAGCTTACGGGACCGTACCGCTCCCTTACACCGACGCGTCTTGCATGAGGCG 1080
 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrArg 380
 1081 CAGCGGCTGCTGGGCTGGTGGCCATGGGAATACCCCGCACCTCATGCGGACCAACCCGC 1140
 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
 1141 TTCCGAGGGTACACCTGCTCCCGAGGACGAGGCTTCTCCCTCTCTGCTCCATCTCTG 1200
 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
 1201 CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTCTGGATGCA 1260
 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
 1261 GATGGACGGTTACAGGAAGCATGAGGGGTTCTCTCCCTCTCTAGGAAAGCGTGTCTGC 1320
 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
 1321 CTGGAAGAGGCTGGGAAAGCGGAGCTTCTCTTCTTCAACCATCTCTACAGGCC 1380
 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480
 1381 TTCTCTCTGAGAGCCGCTGCGCCCGGACACCTCTGAGCTTCAAGCCACCGTCAGTGGC 1440
 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
 1441 CTTTTCACATTCCTCCCGAGGCTTCCAGTGCAGTCCGTCCTCCACTGACCTTCACTCCACC 1500
 501 ThrGlnThrArg 504
 1501 ACGAGACCCAGA 1512

AC AAD60552;
 XX 18-DEC-2003 (first entry)
 XX Human cytochrome p450, 33303 cDNA.
 DE
 XX Human, cytochrome p450; dehydrogenase/reductase; osteomalacia; psoriasis;
 KW lipoxigenase; hydratase; proliferative disorder; haematopoietic disorder;
 KW differentiative disorder; carcinoma; sarcoma; leukaemia; immune disorder;
 KW anorexia nervosa; neural degeneration; muscular disorder; lipid disorder;
 KW multiple sclerosis; encephalomyelitis; metabolic disorder; osteoporosis;
 KW myasthenia gravis; bone metabolism; immunomodulator; obesity; anorectic;
 KW eating disorder; osteodystrophy; arthritis; diabetes; diabetes; rickets;
 KW milk fever; diabetes mellitus; gene; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 1. .1515
 FT /*tag= a
 FT /product= "Human 33303"
 FT /note= "CDS is specifically claimed as SEQ ID NO: 6 in
 FT claim 1 of the specification"
 FT sig_peptide 1. .87
 FT /*tag= b
 FT mat_peptide 88. .1512
 FT /*tag= c
 FT /product= "Mature human 33303"
 XX US2003092658-A1.
 PN 15-MAY-2003.
 XX 20-JUN-2002; 2002US-00175696.
 PF 02-FEB-2001; 2001US-0266140P.
 PR 04-FEB-2002; 2002US-0006766P.
 XX (MEYE/) MEYERS R E.
 PA (GLUC/) GLUCKSMANN M A.
 PA (RUDO/) RUDOLPH-OWEN L A.
 XX Meyers RE, Glucksmann MA, Rudolph-Owen LA;
 PI WPI: 2003-765490/72.
 XX P-PSDB; AAE39880.
 DR New cytochrome P450 family members, dehydrogenase/reductase protein,
 PT lipoxigenase family member and human hydratase useful for treating
 XX cellular proliferative and/or differentiative disorders.
 PS Claim 1; Page 170-171; Opp; English.
 XX The invention relates to isolated cytochrome p450 family members such as
 CC 33312, 33302 and 32579, dehydrogenase/reductase protein such as 21509,
 CC 33770, lipoxigenase family member such as 46638 and human hydratase such
 CC as 50090. The invention is useful as diagnostic targets and agents for
 CC controlling one or more of cellular proliferative and/or differentiative
 CC disorders such as carcinoma, sarcoma or haematopoietic disorders such as
 CC leukaemia, immune disorders such as diabetes mellitus, arthritis,
 CC multiple sclerosis, encephalomyelitis, myasthenia gravis, psoriasis, etc.
 CC or metabolic disorders such as obesity, anorexia nervosa, cachexia, lipid
 CC disorders and diabetes. The invention is useful for controlling disorders
 CC associated with abnormal fatty acid biosynthesis and neural degeneration.
 CC The invention is useful for controlling disorders associated with defects
 CC in fatty acid oxidation or proliferation or muscular disorders and for
 CC treating disorders associated with bone metabolism such as osteoporosis,
 CC osteodystrophy, osteomalacia, rickets, osteitis fibrosa cystica, milk
 CC fever. The present sequence is human cytochrome p450 cDNA
 XX Sequence 1927 BP; 383 A; 625 C; 550 G; 369 T; 0 U; 0 Other;
 SQ Alignment Scores:

Pred. No.:	3.53e-231	Length:	1927
Score:	2615.00	Matches:	504
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-669-693-2 (1-504) x AAD60552 (1-1927)

[illegible]

QY	321	TyrProHisValGlnLysTyrPValArgGluGluLeuAsnArgGluLeuGlyValAlaGlyGln	340
Db	961	TACCCCTCATGTCCAAAGAGTGGTACGTGAGGAGCTGAATCGGAGAGCTGGGGGCTTGGCAG	1020
QY	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
Db	1021	GCACCAAGCCTTAGGGGACCGTACCCTGGCTCCCTTACACCGACGCGGTCTTGTGCATGAGCGC	1080
QY	361	GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrThrArg	380
Db	1081	CAGCGGCTGCTGGCGCTGGTGCCCATGGGAATACCCGACCCCTCATGCGGACCAACCCGC	1140
QY	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
Db	1141	TTCCGAGGGTACACCCCTGCCCCAGGGCACGGAGGTCTTCCCCCTCTTGGCTTCATCCTG	1200
QY	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
Db	1201	CATGACCCCAACATCTTCAAGACACCCAGAGAGTTCACCCAGACCGTTTCTCTGGATGCA	1260
QY	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
Db	1261	GATGGACGGTTACGGAAGCATGAGGCGTCTCTGCCCTTCTCTTAGGAAGCGTGTCTGC	1320
QY	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla	460
Db	1321	CTTGGAGAGGCGCTGGCAAAAGCGGAGCTCTTCTCTTTTACCAACCATCTTACAAGCC	1380
QY	461	PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly	480
Db	1381	TTCTCCCTGGAGAGCCGCTGCCCGCGGACACCTGAGCCTCAAGGCCACCGTCAGTGGC	1440
QY	481	LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr	500
Db	1441	CTTTTCAACATCTCCCGAGCTTCCAGCTTCCAGTGCAAGTCCGTCACCTGACCTTCACTCCACC	1500
QY	501	ThrGlnThrArg	504
Db	1501	ACGCAGACCAGA	1512
RESULT 4			
ABX71208			
ID	ABX71208 standard; cDNA; 2085 bp.		
XX	ABX71208;		
XX	AC		
XX	AC		
DT	05-MAR-2003 (first entry)		
XX	XX		
XX	XX		
XX	Novel human cDNA sequence #433.		
KW	Human; Gene; ss; nervous system disorder; peripheral neuropathy;		
KW	Huntington's disease; amyotrophic lateral sclerosis; haemophilia;		
KW	neurodegenerative disease; Parkinson's disease; Alzheimer's disease;		
KW	autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;		
KW	insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;		
KW	ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;		
KW	fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;		
KW	coagulation disorder; cancer; tumour; inflammatory disease; septic shock;		
KW	Crohn's disease; anaphylaxis; proliferation; chemotactic;		
KW	differentiation; stem cell growth factor; haematopoiesis; chemokinetic;		
KW	haemostatic; antiinflammatory; expressed sequence tag, EST.		
OS	Homo sapiens.		
XX	WO200281731-A2.		
PN	17-OCT-2002.		
XX	XX		
XX	XX		
PF	29-JAN-2002; 2002WO-US001222.		
XX	XX		
PR	30-JAN-2001; 2001US-00774528.		
XX	(HYSE-) HYSEO INC.		

PA (GOOD/) GOODRICH R. W.
XX Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2003-058563/05.
XX Novel polypeptide useful for treating neurodegenerative diseases, myeloid
XX or lymphoid cell disorders, bone disorders, mechanical and traumatic
XX disorders, coagulation disorders, and inflammatory diseases.
XX Claim 1; Page: 612pp; English.
XX This invention relates to the cDNA sequences encoding an isolated novel
XX human polypeptide. The protein encoded by the nucleic acid of the
XX invention is useful for treating central and peripheral nervous system
XX diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic
XX lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,
XX Alzheimer's disease); autoimmune disease (e.g. systemic lupus
XX erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)
XX or myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia)
XX ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,
XX osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head
XX trauma); lung or liver fibrosis; reperfusion injury in various tissues;
XX allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);
XX cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's
XX disease, anaphylaxis). The protein may be used to inhibit the growth,
XX infection or function of infectious agents such as bacteria, fungi,
XX viruses, or to effect bodily characteristics, biorhythms or circadian
XX cycles of rhythms. The protein may also have
XX proliferation/differentiation, stem cell growth factor, haematopoiesis
XX regulation, immune stimulating or suppressing, chemotactic/chemokinetic,
XX haemostatic and thrombolytic, receptor/ligand, and antiinflammatory
XX activities. The cDNA sequences of the invention are useful for expressing
XX recombinant protein for analysis. The present sequence represents a novel
XX human cDNA sequence of the invention, this sequence is an expressed
XX cDNA sequence tag (EST) and was identified using subtractive hybridisation
XX
XX Sequence 2085 BP; 408 A; 691 C; 585 G; 401 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3,936-231 Length: 2085
XX Score: 2615.00 Matches: 504
XX Percent Similarity: 100.00 Conservative: 0
XX Best Local Similarity: 100.00 Mismatches: 0
XX Query Match: 100.00 Indels: 0
XX DB: 7 Gaps: 0
XX
XX US-10-669-693-2 (1-504) x ABX71208 (1-2085)
XX
XX 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20
XX
XX 9 ATGAGCGCACCGCACCTGGGGCGTCTGCTGGCGCTGGCGTCTGCTGCTGACG 68
XX
XX 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40
XX
XX 69 CTGGCGCTGTCCGGGACCGAGGCGCCAGGCGACCTGCCCCCGGGCCACCGCGCTACCA 128
XX
XX 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyValAlaLeuTyrSerGlyLeuMetArgLeu 60
XX
XX 129 CTGCTGGGAACCTCTCCGAGCTACGCGCCGGGGCGCTGTATTACAGGGCTCATCGGCTG 188
XX
XX 61 SerIleLysTyrGlyProValPheThrIleTyrLeuGlyProTrpArgProValVal 80
XX
XX 189 AGTAAGAAGTACGACCGGGTGTTCACCATCTACCTGGGACCCCTGGCGGCTGTGGTGGTC 248
XX
XX 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlnAlaGluGluPheSerGly 100
XX
XX 249 CTGTTGGGAGGAGGCTGTGGGAGGCGCTGGAGGTTCAGGCTGAGGAGTTTCAGCGGC 308
XX
XX 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
XX

Db CGGGAAACCGTAGCGATGCTGGAAGGAGCACTTTTATGGCCATGGGGTTTTCTTCTCCCAAC 368
Qy GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db GGGAGCGGTGGAGCGAGCTGAGGAGTTTACCATGCTTGTCTCGGGAGCTGGCATG 428
Qy GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db GGGAGCGAGAGCGAGGAGCTGATCCAGCGGAGGCCCGGTGTCTGTGGAGACATTC 488
Qy GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
Db CAGGGGACAGAGGACGCCCATTCGATCCCTCCCTGCTGCTGGCCAGCCACCTCCAAC 548
Qy ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db GTAGTCTCTCTCCCTCCCTTTGGCTCCGCTTCTCTATGAGGATAAGAGTTCCAGGCC 608
Qy ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
Db GTGGTCCGGGAGCTGTGGTACCTGCTGGGAGTCACTCCACAGGGGGGTCCAGACCTAC 668
Qy GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisHis 240
Db GAGATGTTCTCTCTGCTGCTGGGCCCTGCGAGGCCCCCAACAGCAGCTCTCTCCACAC 728
Qy ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db GTACACACTTGGCTGCCCTTACAGTCCGGCAGGTGCAGCAGCAGCAGGAGAACCTGGAT 788
Qy AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
Db GCTTCGGGCCCGCGACGCTGACCTTGTGATGCTTCTGCTGAAGATGTCACAGGAGAA 848
Qy GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db CAAACCCAGGACAGAAATTCACCAAGAAATGCTGATGACAGTCAATTATTTCGTG 908
Qy PheAlaGlyThrMetThrValSerThrValGlyTyrThrLeuLeuLeuMetLys 320
Db TTTGCTGGAGCGATGACGCTGAGCACCGCTGCGGTATACCTCTCTCTGCTGATGAAA 968
Qy TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
Db TACCTCTCATGTCAAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGTGGCCAG 1028
Qy AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db GCACCAAGCTTAGGGACCGTACCCGCTCCCTTACACGACCGCGTTCGCTGATGAGCG 1088
Qy GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
Db CAGCGCTGCTGGCGTGGTGGCCATGGGAATACCCCGCACCTCATCGGAGCACCCCGC 1148
Qy PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db TTCGAGGGGTACACCTGCCCCCAGGGCAGCGAGTCTTCCCTCTCTCTTGGCTCCATCCTG 1208
Qy HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db CATGACCCCCAACATCTTCAAGCACCCAGAGAGTTCAACCCAGACCGTTCCTGGATGCA 1268
Qy AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db GATGGACGGTTCAGGAGCATGAGGGGTCTCTGCCCTTCTCTTAGGAAAGCGTGTCTGC 1328
Qy LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAla 460
Db CTTGGAGAGGCGCTGGCAAAAGCGGAGCTCTTCTCTTCTTCTTACCAACCATCTCTACAGCC 1388
Qy PheSerLeuLeuSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480
Db TTCTCCCTGGAGAGCGCGCTGCCCGCGCACACCTGAGCTCAAGCCACCGCTCAGTGGC 1448

QY 481 LeuPheAenileProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
 DB 1449 CTTTTCACATTCCTCCAGCCCTCCAGCTGCAGTCCGTCCTCCACTGACCTTCACTCCACC 1508

QY 501 ThrGlnThrArg 504
 DB 1509 ACGCAGACCAGA 1520

RESULT 5
 ABQ72550
 ID ABQ72550 standard; cDNA; 2604 BP.
 AC ABQ72550;
 XX
 DT 03-SEP-2002 (first entry)
 XX
 DE Human MDDT encoding cDNA SEQ ID NO 102.
 XX
 KW Human; MDDT; disease detection and treatment molecule polynucleotide;
 KW proliferative disorder; hepatitis; psoriasis; cancer; AIDS;
 KW autoimmune disorder; inflammatory disorder; allergy; multiple sclerosis;
 KW rheumatoid arthritis; transgenic; gene therapy; antiarteriosclerotic;
 KW hepatotropic; antiinflammatory; antipsoriatic; cytostatic; anti-HIV;
 KW antiallergic; antianaemic; antiaethmatic; antiatherosclerotic; antigout;
 KW neuroprotective; antirheumatic; antiarthritic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX W0200240715-A2.
 PN
 XX
 PD 23-MAY-2002.
 XX
 XX
 PF 06-SEP-2001; 2001WO-US027628.
 XX
 XX
 PR 05-SEP-2000; 2000US-0229747P.
 PR 05-SEP-2000; 2000US-0229748P.
 PR 05-SEP-2000; 2000US-0229749P.
 PR 05-SEP-2000; 2000US-0229750P.
 PR 05-SEP-2000; 2000US-0229751P.
 PR 05-SEP-2000; 2000US-0230583P.
 PR 05-SEP-2000; 2000US-0230584P.
 PR 06-SEP-2000; 2000US-0230505P.
 PR 06-SEP-2000; 2000US-0230514P.
 PR 06-SEP-2000; 2000US-0230515P.
 PR 06-SEP-2000; 2000US-0230517P.
 PR 06-SEP-2000; 2000US-0230518P.
 PR 06-SEP-2000; 2000US-0230519P.
 PR 06-SEP-2000; 2000US-0230595P.
 PR 06-SEP-2000; 2000US-0230597P.
 PR 06-SEP-2000; 2000US-0230598P.
 PR 06-SEP-2000; 2000US-0230599P.
 PR 06-SEP-2000; 2000US-0230610P.
 PR 06-SEP-2000; 2000US-0230865P.
 PR 06-SEP-2000; 2000US-0230988P.
 PR 06-SEP-2000; 2000US-0230989P.
 PR 07-SEP-2000; 2000US-0230951P.
 PR 07-SEP-2000; 2000US-0231163P.
 PR 07-SEP-2000; 2000US-0231167P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Jackson S, Lincoln SE, Altus CM, Dufour GE, Chalup MS;
 PI Hillman JL, Jones AL, Wright RJ, Gietzen D, Liu TF, Yap PE;
 PI Dahl CR, Momiyama MG, Bradley DL, Rohatgi SD, Harris B;
 PI Roseberry AM, Gerstin EH, Peralta CH, David MH, Panzer SR,
 PI Daffo A, Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
 XX WPI; 2002-527544/56.
 DR P-PSDB; ABP51332.
 DR
 XX Novel human disease detection and treatment polypeptide, useful in
 PT diagnosis, prevention or treatment of cell proliferative disorders e.g.
 PT arteriosclerosis, cirrhosis and an autoimmune/inflammatory disorder e.g.

PT AIDS.
 XX
 PS Claim 1; Page 364-365; 618pp; English.
 XX
 CC The invention relates to an isolated human disease detection and
 CC treatment (MDDT) polypeptide (I) selected from a polypeptide having a
 CC sequence selected from 254 sequences (ABP51231-ABP51484) given in the
 CC specification, a naturally occurring polypeptide comprising a sequence
 CC having at least 90% identity to (I) or a biologically active or
 CC immunogenic fragment of (I). (I) is useful for screening a compound for
 CC effectiveness as an agonist or antagonist, for screening a compound that
 CC specifically binds (I) or modulates the activity of (I), and for
 CC preparing a polyclonal or monoclonal antibody by hybridoma technology.
 CC Nucleic acids (II) (ABQ72449-ABQ72700) encoding (I) are useful for
 CC screening a compound for effectiveness in altering expression of a target
 CC polynucleotide comprising. Oligonucleotides and antibodies are useful for
 CC detecting MDDT in a sample or for assessing toxicity of a test compound,
 CC in a diagnostic test for a condition or a disease associated with the
 CC expression of MDDT in a biological sample, for detecting (I) in a sample,
 CC and for purifying (I) from a sample. A composition comprising (I), an
 CC agonist or antagonist is useful for treating a disease or condition
 CC associated with decreased or increased expression of functional MDDT. (I)
 CC or (II) are useful for diagnosing, treating or preventing disorders
 CC associated with aberrant expression of MDDT, where the disorders are
 CC selected from a cell proliferative disorder such as arteriosclerosis,
 CC cirrhosis, hepatitis, psoriasis, and cancer and an
 CC autoimmune/inflammatory disorder such as AIDS, Addison's disease,
 CC allergy, anaemia, asthma, atherosclerosis, gout, multiple sclerosis or
 CC rheumatoid arthritis. (II) are useful for creating knockin humanised
 CC animals or transgenic animals to model human diseases, in somatic or
 CC germline gene therapy, to generate a transcript image of a tissue or cell
 CC type, for detecting differences in the chromosomal location due to
 CC translocation or inversion among normal, carrier or affected individuals
 CC and as hybridisation probes for mapping naturally occurring genomic
 CC sequences
 XX
 SQ Sequence 2604 BP; 533 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5-32e-231 Length: 2604
 Score: 2615.00 Matches: 504
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-10-669-693-2 (1-504) x ABQ72550 (1-2604)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20
 DB 40 ATGGAGCGCAGCCGACCTGGCGCTGCTGGCGCTGGCGCTGGCGCTGGCGCTGACG 99
 QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40
 DB 100 CTGGCGCTGTCCGGACCCAGGCCACCTGCCCCCGGGCCACCGCGCTACCA 159
 QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
 DB 160 CTGCTGGGAACCTCTCTGAGCTACGGCCCGGGCGCTGTATTGAGGCTCATGGCGCTG 219
 QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValValVal 80
 DB 220 AGTAAGAAGTACGACCGCGTGTCCACCTACCTGGGACCCCTGGCGGCTGGTGGTGC 279
 QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
 DB 280 CTGGTTGGCGCAGAGGCTGTCCGGAGGCCCTGGGAGGTTCAGGCTCAGAGTTCAGCGC 339
 QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
 DB 340 CGGGGAACCGTAGCGATCTGGAGGGGACTTTTGATGGCCATGGGGTTTCTTCTCCAAC 399
 QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140

1480 CTTTTCACATTCCTCCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1539

Db 400 GGGGAGCGGTGGAGCGAGCTGAGGAAGTTTACCATGCTTGTCTCGGGAGCCTGGGCAATG 459
 Qy 141 GLVLSARGGLUGLYGLUGLULEGLNAlaGLUAlaAArgCysLeuValGluThrPhe 160
 Db 460 GGGAGCGAGGAGGAGGAGCTGATCAGCGGAGGCGCGGTGTCTGTGGAGACATTC 519
 Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180
 Db 520 CAGGGGACAGAGGAGCGCCATTCATCCCTCCCTGCTGGCCAGCGCCACCTCCAAAC 579
 Qy 191 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
 Db 580 GTAGTCTCTCCCTCTCTTTGGCTCCGCTTCTCTATGAGGATAGAGATTCCAGGCC 639
 Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyr 220
 Db 640 GTGGTCCGGGAGCTGGTGGTACCTGCTGGAGTCAGTCCAGGGGGTTCAGACCTAC 699
 Qy 221 GluMetPheSerTyrPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
 Db 700 GAGATGTTCTCTGCTGCTGGGCGCTCCAGGCGCCCAACAGCAGCTCTCCACCCAC 759
 Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
 Db 760 GTCAGCACCTTGGCTGCTTTCACAGTCCGCGAGGTCAGCAGCACCGGGGAACTGGAT 819
 Qy 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
 Db 820 GTTTCGGGCGCGAGCTGACCTTGTGATGCTTCTCTGCTGAAGATGGCAGAGGAA 879
 Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
 Db 880 CAAACCCAGGACAGATTACCAACAAGACATGCTGATCAGCTCATTTATTGTGCG 939
 Qy 301 PheAlaGlyThrMetThrValSerThrValGlyTyrThrLeuLeuLeuMetLys 320
 Db 940 TTTGTGGGACGATGACGTCAGCACCGTCGCTATACCTCTCTCTGATGATA 999
 Qy 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
 Db 1000 TACCCTCATGTCACAAAGTGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGTGGCCAG 1059
 Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
 Db 1060 GCACCAAGCTAGGAGCGCTACCCGCTCCCTTACACCGACCGGTTCGTGATGAGCGG 1119
 Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrThrArg 380
 Db 1120 CAGCGGCTGTGGCGTGGTGGCCATGGGATATACCCGCACTCATCGGACCAACCCG 1179
 Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
 Db 1180 TTCGAGGTTACACCTGCCCCAGGCGACGAGGTCTTCCCTCTCTGCTGCTGCTGCTG 1239
 Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
 Db 1240 CATGACCCCAACATCTTCAAGACCCCAAGAGAGTTCACCCAGACCGTTTCTTGGATGCA 1299
 Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
 Db 1300 GATGGAGGTTTCAAGACATGAGCGGTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTT 1359
 Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
 Db 1360 CTTGGAGGCGCTGGCAAAAGCGAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1419
 Qy 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480
 Db 1420 TTTCTCTGGAGACCGGTGCGCGGACACCTGAGCTTCAAGCCACCGTCACTGCTGGC 1479
 Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500

Db 1480 CTTTTCACATTCCTCCAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1539
 Qy 501 ThrGlnThrArg 504
 Db 1540 ACGCAGACCAGA 1551
 RESULT 6
 ID AAD24671 standard; cDNA; 2620 BP.
 AC AAD24671;
 XX 12-MAR-2002 (first entry)
 DT Human drug metabolising enzyme (DME)-6 cDNA.
 DE Human; drug metabolising enzyme; gene therapy; autoimmune disorder;
 KW inflammatory disorder; acquired immune deficiency syndrome; infection;
 KW AIDS; adult respiratory distress syndrome; cell proliferative disorder;
 KW allergy; anaemia; conjunctivitis; actinic keratosis; arteriosclerosis;
 KW cancer; endocrine disorder; hypothalamus disorder; pituitary disorder;
 KW gastrointestinal disorder; metabolic disorder; developmental disorder;
 KW liver disorder; iritis; cystic fibrosis; Addison's disease; retinitis;
 KW goitre; diabetes; dysphagia; gastric carcinoma; anorexia; eye disorder;
 KW DME-6; ss.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 41..1555
 FT /tag= a
 FT /product= "Human DME-6 protein"
 FT sig_peptide 41..124
 FT /tag= b
 FT mat_peptide 125..1552
 FT /tag= c
 FT /product= "Human mature DME-6 protein"
 XX WO200179468-A2.
 PD 25-OCT-2001.
 XX 12-APR-2001; 2001WO-US011869.
 XX 13-APR-2000; 2000US-0197590P.
 PR 19-APR-2000; 2000US-0198403P.
 PR 28-APR-2000; 2000US-0200185P.
 PR 05-MAY-2000; 2000US-020234P.
 PR 11-MAY-2000; 2000US-0203509P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Policky JL, Hafalia A, Burford N, Ring HZ, Lal P, Tribouley CM;
 PI Yao MG, Yue H, Tang YT, Patterson C, Das D, Sanjanwala MS;
 PI Gandhi AP, Reddy R, Khan FA, Baughn MR, Ramkumar J, Griffin JA;
 PI Au-Young J;
 XX WPI; 2002-066363/09.
 DR P-PSDB; AAE15439.
 XX Novel isolated human drug metabolizing enzymes referred as DME 1-10,
 PT useful for diagnosing, treating, or preventing disorders associated with
 PT aberrant expression of DME such as allergy, anemia, asthma, infertility.
 XX Claim 5; Page 139-140; 143pp; English.
 PS The invention relates to human drug metabolising enzymes referred as DME
 CC and nucleic acid molecules encoding such enzymes. Polynucleotides of the
 CC invention are useful for assessing toxicity of test compounds and in gene
 CC therapy. Sequences of the invention are useful in the diagnosis,
 CC prevention and treatment of autoimmune/inflammatory disorders such as
 CC acquired immune deficiency syndrome (AIDS), adult respiratory distress
 CC syndrome, allergies, anaemia, atherosclerosis, asthma, autoimmune

CC haemolytic anaemia, contact dermatitis, Crohn's disease,
 CC glomerulonephritis, Goodpasture's syndrome, gout, Grave's disease,
 CC Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis,
 CC osteoarthritis, osteoporosis, psoriasis, systemic lupus erythematosus,
 CC rheumatoid arthritis, scleroderma, ulcerative colitis, uveitis, viral,
 CC bacterial, fungal, parasitic, protozoal, helminthic infections; cell
 CC proliferative disorders such as actinic keratosis, arteriosclerosis,
 CC atherosclerosis, Duchenne and Becker dystrophy, hepatitis, Cushing's
 CC syndrome, cancers, myelodysplastic syndrome, epilepsy; endocrine
 CC disorders such as disorders of the hypothalamus and pituitary resulting
 CC from lesions such as primary brain tumours, adenomas, infarction
 CC associated with pregnancy, aneurysms, vascular malformations; eye
 CC disorders such as conjunctivitis, iritis, retinitis, glaucoma, pigmentosa
 CC ; metabolic disorders such as Addison's disease, cystic fibrosis,
 CC diabetes, gout, glycogen storage diseases, hypercholesterolaemia,
 CC hyperthyroidism, hypoglycaemia, lipid myopathies, Menkes syndrome,
 CC mannosidosis, obesity; gastrointestinal disorders such as dysphagia,
 CC gastric carcinoma, anorexia, nausea, gastroenteritis,
 CC hyperbilirubinaemia, emesis, cirrhosis, diarrhoea, jaundice, Reye's
 CC syndrome, peliosis hepatitis, hepatic vein thrombosis and developmental
 CC disorders. The present sequence is human DME-6 protein cDNA
 XX

SQ Sequence 2620 BP; 549 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.36e-231 Length: 2620
 Score: 2615.00 Matches: 504
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-669-693-2 (1-504) x AAD24671 (1-2620)

Qy 1 MetGluAlaThrGlyThrTrpAlaLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThr 20
 Db 41 ATGAGGCGACCGGACCTGGGCGCTCTGTGGCGTGGCGTCTCTGTCTGCTGACG 100
 Qy 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProGlyProLeuPro 40
 Db 101 CTGGCGTGTCTGGGACCGAGGCGGAGGCCACTGTGCCCCCGGGCCCGCGCTACCA 160
 Qy 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
 Db 161 CTCTGGGAACCTCTCGAGTACGCGCGGCGCGCTGTATTACGGCTCATGGCGCTG 220
 Qy 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValVal 80
 Db 221 AGTAAGAAGTACGAGCCGGTGTTCACCATCTACTCGGACCTGTGGCGCTGTGGTGC 280
 Qy 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluPheSerGly 100
 Db 281 CTGGTTGGGCGAGGCTGTGGGAGGCGCTGGAGGTCAGGCTGAGGAGTTTCAGCGGC 340
 Qy 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
 Db 341 CGGGGAACCTGATCGATGCTGGAAGGACTTTTGTATGGCCATGGGTTTCTTCTCCAC 400
 Qy 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
 Db 401 GGGGAGCGGTGGAGGCGAGCTGAGGAAGTTTACCATGCTTGTCTGTGGGACCTGGGCATG 460
 Qy 141 GlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160
 Db 461 GCGAAGCGAGAGGCGAGGCTGATCCAGGCGGAGGCGCGGTGTCTGTGGAGACATTC 520
 Qy 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
 Db 521 CAGGGGACAGAGGACGCGCAATTCATGATCCCTCCCTGCTGTGCCCGAGCCACTCCAC 580
 Qy 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
 Db 581 GTAGTCTGCTCTCTCTTTGGCGCTCGCGCTTCTCTATGAGGATAAGGAGTTCCAGGCC 640

Qy 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
 Db 641 GTGGTCCGGGACGTGGTGGTACCTGTCTGGGAGTCAGTCCACAGGGGGGTGACACTAC 700
 Qy 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
 Db 701 GAGATGTTCTCTGTGGTCTCTGGGCGCTTGGCAGGCCCTCCACAGCAGCTCTCTCCACCAC 760
 Qy 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
 Db 761 GTCAGCACCTTGGTGGCTTTCAGTCCGGCAGGTGCAGCAGCACAGGGGAACCTGGAT 820
 Qy 261 AlaSerGlyProAlaArgAspLeuValAlaPheLeuLeuLysMetAlaGlnGluGlu 280
 Db 821 GCTTGGGCGCGGACGTGATGCTTGTGATGCTTCTGTGCTGAAGATGCGCAGAGGANA 880
 Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
 Db 881 CAAACCCAGGCGACAGAAATTCACCAACAAGAACATGCTGTGATGACAGTCAATTTATTTGCTG 940
 Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuLeuMetLys 320
 Db 941 TTTGCTGGGACGATGACGGTCAGCACCGTTCGGCTATACCTCTCTCTCTCTGATGAAA 1000
 Qy 321 TyrProHisValGlnLysTrpValArgGluLeuAsnArgGluLeuGlyAlaGlyGln 340
 Db 1001 TACCTCATGTCTCAAAAAGTGGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAG 1060
 Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
 Db 1061 GCACCAAGCCTTAGGGGACCGGTACCCGCTCTTACACCGACGCGTCTCATGAGCGC 1120
 Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
 Db 1121 CAGCGGCTGTGGCGTGTGGTCCCATGGGAATACCCCGCACTCATGCGGACCAACCGCG 1180
 Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
 Db 1181 TTCGAGGGTACACCTGTGCCCGGAGGCGAGGCTCTTCCCTCTCTCTCTCTCTCTCTCT 1240
 Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
 Db 1241 CATGACCCCAACATCTTCAAGACCCAGAGAGTTCACCCAGACCGCTTCTCTGGATGCA 1300
 Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
 Db 1301 GATGGACGGTTCAGGAAGCATGAGGGCTTCTGCCCTTCTCTTAGGAGCGGTGTCTGC 1360
 Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrIleLeuGlnAla 460
 Db 1361 CTGGAGAGGGCTGGCAAAAGCGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1420
 Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
 Db 1421 TTCTCTCTGGAGGCGCGTGGCGCGGACACCTCAGGCTCAAGGCCCAACCGTCAGTGGC 1480
 Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
 Db 1481 CTTTTCACATCTCCCCAGCCTTCAGCTGCAGTCCGTCCTCCCTCTCTCTCTCTCTCTCTCT 1540
 Qy 501 ThrGlnThrArg 504
 Db 1541 ACGCAGACACAGA 1552

RESULT 7

AAS21275

ID AAS21275 standard; cDNA; 2623 BP.

XX

AC AAS21275;

XX

DT 24-OCT-2001 (first entry)

XX

CC polypeptides, to link bioactive molecules to cells expressing PRO
CC polypeptides, to modulate biological activities of cells expressing PRO
CC polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample. Some
CC of the 275 sequences are also useful to stimulate the release of tumour
CC necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or
CC differentiation of chondrocytes, the proliferation or gene expression in
CC pericyte cells, the release of proteoglycans from cartilage, the
CC proliferation of inner ear utricular supporting cells or of T-
CC lymphocytes, the release of a cytokine from peripheral blood monocytes
CC (PBMCs), or the proliferation of endothelial cells. Some of the PRO
CC polypeptides may modulate glucose or free fatty acid uptake by skeletal
CC muscle cells or by adipocytes; or inhibit binding of A-peptide to factor
CC VIIA. The PRO polypeptides can be used in assays to identify molecules
CC involved in binding interactions. The polynucleotides encoding PRO
CC polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy
XX SQ Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,37e-231 Length: 2623
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-669-693-2 (1-504) x AAS21275 (1-2623)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuThr 20
Db 41 ATGGAGGCGACCGGCACCTGGCGCTGCTGCTGGCGCTGGCGCTGCTGCTGGCG 100
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro 40
Db 101 CTGGCGCTGTCCGGGACCGAGGCGCCGAGGCGACCTGCCCGCGGCGCCGCGCTACCA 160
QY 41 LeuLeuGlyAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
Db 161 CTGGCGCTGTCCGGGACCGAGGCGCCGAGGCGACCTGCCCGCGGCGCCGCGCTACCA 220
QY 61 SerLysLysThrGlyProValPheThrIleTyrLeuGlyProThrProValValVal 80
Db 221 AGTAAGAAGTACGAGCGCGGTGTTACCACTACCTGGGACCCCTGGCGGCGCTGTGGTGC 280
QY 81 LeuValGlyGlnGlnAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
Db 281 CTGGTTGGGCGAGGAGCTGTGGGAGGCGGCTGGGAGGCTGAGGAGTTCAGGCGGC 340
QY 101 AtgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 341 CGGGGAACCGTAGCATGCTGGGAAGGACCTTTTATGGCCATGGGTTTTCTTCTCCAAC 400
QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db 401 GGGGAGCGGTGGAGGCGAGCTGAGGAAATTTTACCATGCTTCTGCGGGACCTGGGCGATG 460
QY 141 GlyIleArgGluGlyGluGluLeuLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db 461 GGGGAACGAGAGCGAGGAGCTGATCCAGCGGAGGCGGCTGTCTGTGGGAGACATTC 520
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180
Db 521 CAGGGACAGAGGACGCCCATTCGATCCCTGCTGCTGGCGCCAGGCGACCTCCCAAC 580
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db 581 GTAGTCTGCTCCCTCTCTTTGGCCCTCCCTTATGAGGATAGGAGTTCAGGCC 640
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220

DE Human cDNA sequence encoding for PRO1906 polypeptide.
XX Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;
KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;
KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;
KW A-peptide; factor VIIA; gene therapy; ss.
XX Homo sapiens.
OS WO200140466-A2.
PN 07-JUN-2001.
XX 01-DEC-2000; 2000WO-US032678.
XX 01-DEC-1999; 99WO-US028301.
XX 01-DEC-1999; 99WO-US028634.
XX 02-DEC-1999; 99WO-US028551.
XX 02-DEC-1999; 99WO-US028584.
XX 02-DEC-1999; 99WO-US028565.
XX 09-DEC-1999; 99US-0170262P.
XX 16-DEC-1999; 99WO-US030095.
XX 20-DEC-1999; 99WO-US030911.
XX 20-DEC-1999; 99WO-US030999.
XX 30-DEC-1999; 99WO-US031243.
XX 30-DEC-1999; 99WO-US031274.
XX 05-JAN-2000; 2000WO-US000219.
XX 06-JAN-2000; 2000WO-US000277.
XX 06-JAN-2000; 2000WO-US00376.
XX 11-FEB-2000; 2000WO-US003565.
XX 18-FEB-2000; 2000WO-US004341.
XX 18-FEB-2000; 2000WO-US004342.
XX 22-FEB-2000; 2000WO-US004414.
XX 24-FEB-2000; 2000WO-US004914.
XX 24-FEB-2000; 2000WO-US005004.
XX 01-MAR-2000; 2000WO-US005601.
XX 02-MAR-2000; 2000WO-US005841.
XX 03-MAR-2000; 2000US-0187202P.
XX 10-MAR-2000; 2000WO-US006319.
XX 15-MAR-2000; 2000WO-US006884.
XX 20-MAR-2000; 2000WO-US007377.
XX 21-MAR-2000; 2000WO-US007532.
XX 30-MAR-2000; 2000WO-US008439.
XX 17-MAY-2000; 2000WO-US013705.
XX 22-MAY-2000; 2000WO-US014042.
XX 30-MAY-2000; 2000WO-US014941.
XX 02-JUN-2000; 2000WO-US015264.
XX 05-JUN-2000; 2000US-0209832P.
XX 28-JUL-2000; 2000WO-US020710.
XX 11-AUG-2000; 2000WO-US022031.
XX 23-AUG-2000; 2000WO-US023522.
XX 24-AUG-2000; 2000WO-US023328.
XX 08-NOV-2000; 2000WO-US030952.
XX 10-NOV-2000; 2000WO-US030873.
XX (GETH) GENENTECH INC.
PA Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2001-408281/43.
DR P-PSDB; AAU12203.
XX Isolated , secretory and transmembrane PRO polypeptide used to detect
XX other PRO polypeptides, link bioactive molecules to cells expressing PRO
PT polypeptides, and detect the presence of mammalian tumors e.g. lung,
PT breast, prostate, cervical.
XX Claim 3; Fig 63; 813pp; English.
XX AAS21244-AAS21518 encode for novel human secretory and transmembrane PRO
CC polypeptides. The PRO polypeptides are useful to detect other PRO
CC

Db 641 GTGGTCCGGGAGCTGGTGGTACCTCTGGAGTCAGCTCCAGGGGGTTCAGACCTAC 700
 QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisIysGlnLeuLeuHisHis 240
 Db 701 GAGATGTTCTCTGGTCTCTGGCCCTTCAGGGCCCTCCAGGCCCTCCCAAGCAGCTCTCCACAC 760
 QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
 Db 761 GTCAGCAGCTTGGCTGCTTCACAGTCCGGCAGGTGTCAGCAGCACCAGGGGAACCTGGAT 820
 QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 280
 Db 821 GCTTCGGGCCCGCAGCTGAGCTTCGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
 QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
 Db 881 CAAAACCCAGCAGCAGATTCACCAACAGAACATGCTGATGACAGTCAITTAITTTGCTG 940
 QY 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
 Db 941 TTTTCTGGGACGATGACGGTCAGCACACGGTCGGCTATACCTCTCTCTCTGATGAAA 1000
 QY 321 TyrProHisValGlnLysTyrValArgGlnGluLeuAsnArgGluLeuGlyAlaGlyGln 340
 Db 1001 TACCTCATGTCCAAAGTGGTACGTGAGGAGCTGAATCGGGAGCTGGGGCTGGCCAG 1060
 QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
 Db 1061 GCACCAAGCCTAGGGACCGTACCCGCTCTCTTACACCGACGGGTTCGATGAGCGC 1120
 QY 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
 Db 1121 CAGCGGCTGTGGCGCTGGTGGCCATGGGAATACCCGACCCCTCATGGGACCAACCCCGC 1180
 QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
 Db 1181 TTCGAGGGTACACCTTCAGGACCCGAGGAGGAGTCTTCCCTCTCTTGGCTCCATCCTG 1240
 QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
 Db 1241 CATGACCCCAACATCTCAAGCACCCAGAGAGTTCAACCCAGACCCGTTTCTCGATGCA 1300
 QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
 Db 1301 GATGACGGTTCAGGAAGCATGAGCGTCTTCCCTCTCTTAGGAAGCGTGTCTGC 1360
 QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAla 460
 Db 1361 CTTGGAGAGGCGCTGGCAAGCGAGCTCTCTCTCTTCCACCACTTCTTACCAAGCC 1420
 QY 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480
 Db 1421 TTTCTCTGGAGAGCCGTCGCGCCGCGGACACCTGAGCCTCAAGCCCACTGATGGC 1480
 QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
 Db 1481 CTTTTCACATTCCTCCAGCTTCCAGTCCAGTCCAGTCCGTCAGTCCGCTTCTACTCCACC 1540
 QY 501 ThrGlnThrArg 504
 Db 1541 ACGCAGACCAGA 1552
 RESULT 8
 ACD23884
 ID ACD23884 standard; cDNA; 2623 BP.
 XX
 AC ACD23884;
 XX
 DT 26-AUG-2003 (first entry)
 DE Novel human secreted and transmembrane protein PRO1906 cDNA.
 XX
 KW Human; secreted and transmembrane protein; PRO; antiinflammatory;

KW antiarteriosclerotic; cardiant; anti-infertility; anti-HIV; cytostatic;
 KW antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;
 KW TNF-alpha release; cell proliferation; cell differentiation;
 KW gene expression modulator; proteoglycan release; cytokine release;
 KW tumour; inflammatory disease; organ failure; atherosclerosis;
 KW cardiac injury; infertility; birth defect; premature aging; AIDS;
 KW acquired immunodeficiency syndrome; cancer; diabetic complication;
 KW chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;
 KW bioreactor; tissue typing; gene; ss.
 XX Homo sapiens.
 OS
 XX
 XX US2003032156-A1.
 PD 13-FEB-2003.
 XX
 PF 06-MAY-2002; 2002US-00140474.
 XX
 PR 31-MAR-1997; 97WO-US005230.
 PR 12-JUN-1998; 98WO-US012456.
 PR 14-JUL-1998; 98WO-US014552.
 PR 28-AUG-1998; 98WO-US017888.
 PR 10-SEP-1998; 98WO-US018824.
 PR 14-SEP-1998; 98WO-US019094.
 PR 14-SEP-1998; 98WO-US019177.
 PR 14-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98WO-US019437.
 PR 07-OCT-1998; 98WO-US021141.
 PR 29-OCT-1998; 98WO-US022991.
 PR 29-OCT-1998; 98WO-US022992.
 PR 20-NOV-1998; 98WO-US024855.
 PR 01-DEC-1998; 98WO-US025108.
 PR 05-JAN-1999; 99WO-US000106.
 PR 08-MAR-1999; 99WO-US005028.
 PR 10-MAR-1999; 99WO-US005190.
 PR 20-APR-1999; 99WO-US008615.
 PR 14-MAY-1999; 99WO-US010733.
 PR 02-JUN-1999; 99WO-US012252.
 PR 01-SEP-1999; 99WO-US020111.
 PR 08-SEP-1999; 99WO-US020594.
 PR 13-SEP-1999; 99WO-US020944.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.
 PR 05-OCT-1999; 99WO-US023089.
 PR 29-NOV-1999; 99WO-US028214.
 PR 30-NOV-1999; 99WO-US028313.
 PR 30-NOV-1999; 99WO-US028409.
 PR 01-DEC-1999; 99WO-US028301.
 PR 01-DEC-1999; 99WO-US028634.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028564.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 20-DEC-1999; 99WO-US030999.
 PR 22-DEC-1999; 99WO-US030720.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 11-FEB-2000; 2000WO-US003376.
 PR 18-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 22-FEB-2000; 2000WO-US004342.
 PR 24-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 01-MAR-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005746.
 PR 10-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006319.
 PR 2000WO-US006884.

20-MAR-2000; 2000WO-US007377.
21-MAR-2000; 2000WO-US007532.
30-MAR-2000; 2000WO-US008439.
17-MAY-2000; 2000WO-US013705.
22-MAY-2000; 2000WO-US014042.
30-MAY-2000; 2000WO-US014941.
02-JUN-2000; 2000WO-US015264.
28-JUL-2000; 2000WO-US020710.
11-AUG-2000; 2000WO-US022031.
23-AUG-2000; 2000WO-US023522.
24-AUG-2000; 2000WO-US023328.
08-NOV-2000; 2000WO-US030952.
10-NOV-2000; 2000WO-US030873.
01-DEC-2000; 2000WO-US032678.
20-DEC-2000; 2000US-00747259.
20-DEC-2000; 2000WO-US034956.
28-FEB-2001; 2001US-00796498.
28-FEB-2001; 2001WO-US006520.
01-MAR-2001; 2001WO-US006666.
09-MAR-2001; 2001US-00802706.
14-MAR-2001; 2001US-00808689.
22-MAR-2001; 2001US-00816744.
05-APR-2001; 2001US-00828366.
10-MAY-2001; 2001US-00854208.
10-MAY-2001; 2001US-00854280.
18-MAY-2001; 2001US-00860216.
25-MAY-2001; 2001US-00866028.
25-MAY-2001; 2001US-00866034.
25-MAY-2001; 2001WO-US017092.
01-JUN-2001; 2001US-00872035.
01-JUN-2001; 2001WO-US017800.
05-JUN-2001; 2001US-00874503.
14-JUN-2001; 2001US-00882636.
19-JUN-2001; 2001US-00886342.
20-JUN-2001; 2001WO-US019692.
21-JUN-2001; 2001US-00887879.
22-JUN-2001; 2001WO-US020116.
29-JUN-2001; 2001WO-US021066.
09-JUL-2001; 2001WO-US021735.
18-JUL-2001; 2001US-00908827.
06-AUG-2001; 2001US-00924419.
19-AUG-2001; 2001US-00927796.
16-AUG-2001; 2001US-00931836.
19-DEC-2001; 2001US-00028072.
(GETH) GENENTECH INC.
XX Baker KP, Beresini M, Deforge L, Deenoyers L, Filvaroff E, Gao W;
PI Gerritsen MB, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI: 2003-341980/32.
DR P-PSDB; ABO17647.
XX
XX
XX New secreted and transmembrane PRO nucleic acids, for treating
PT inflammation, organ failure, atherosclerosis, cardiac injury,
PT infertility, birth defects, premature aging, acquired immunodeficiency
PT syndrome (AIDS), or cancer.
XX
XX
XX Claim 2; Fig 63; 660pp; English.
XX
XX The invention describes an isolated nucleic acid (I) comprising, or which
CC has 80 % sequence identity to, or the full-length coding sequence of, one
CC of 275 nucleotide sequences, and which encodes a corresponding
CC polypeptide selected from 275 amino acid sequences, where all sequences
CC are given in the specification. The polypeptide encoded by (I) is used to
CC detect PRO polypeptides, link a bioactive molecule to a cell expressing a
CC PRO polypeptide, modulate a biological activity of a cell, stimulate the
CC release of tumour necrosis factor (TNF)-alpha from human blood, modulate
CC the uptake of glucose or free fatty acid by cells, stimulate or inhibit
CC the proliferation or differentiation of cells or gene expression,
CC stimulate the release of proteoglycans, stimulate the release of cytokine
CC from peripheral blood mononuclear cells, inhibit the binding of A-peptide

CC to factor VIIA, or detect the presence of tumour in a mammal. The nucleic
CC acid and polypeptide encoded by it, are useful for treating inflammatory
CC diseases, organ failure, atherosclerosis, cardiac injury, infertility,
CC birth defects, premature aging, acquired immunodeficiency syndrome
CC (AIDS), cancer, or diabetic complications. The nucleic acid is useful as
CC hybridisation probes, in chromosome and gene mapping, and in generating
CC antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,
CC diagnostics, biosensors or bioreactors. Both are useful in tissue typing.
CC This sequence encodes a novel human secreted and transmembrane PRO
CC polypeptide
XX
SQ Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5,37e-231 Length: 2623
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0
US-10-669-693-2 (1-504) x ACD23884 (1-2623)
QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuThr 20
Db 41 ATGGAGGCGACCGGCACCTGCTGCTGGCGCTGGCGCTGCTGCTGCTGCTGCG 100
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40
Db 101 CTGGCGCTGCTCGGAGCACCGAGGCCACCTGCGCCCGCGGCCCGCGGTACCA 160
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuLeuLeuLeuLeuMetArgLeu 60
Db 161 CTGCTGGGAAACCTCTCTGAGCTACGGCCCGCGCGCTGCTATTACGGGCTCATCGCGCTG 220
QY 61 SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValVal 80
Db 221 AGTAAGAGTACGACCGGTGTTACCATCTACCTGGGACCTCGCGCGCTGGTGGTGC 280
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
Db 281 CTGGTTGGCGAGGAGCTGTGGGAGGCCCTGGAGGCTCAGGCTCAGGAGTTCCGGCG 340
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 341 CGGGGNAACGTAGCGATGCTGGAGGGACTTTTGTATGGCCATGGGCTTTCTTCTCCAAC 400
QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db 401 GGGGAGCGGTGGAGGAGCTGAGGAAAGTTTACCATCTGCTCTCGGGACCTGGGCGATG 460
QY 141 GlyLysArgGluGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db 461 GGGAGCGAGAGGCGGAGGAGCTGATCCAGCGGAGGCCCGCGTGTCTGGTGGAGACATTC 520
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsn 180
Db 521 CAGGGGACAGAGGACGCCCATTCGATCCCTCTCTGCTGCGCCCGGCGACCTCCAAC 580
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db 581 GTAGTCTGCTCCCTCTTGGCCCTTCTGCTTATGAGGATAGGAGTTCCAGGCC 640
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
Db 641 GTGCTCGGCGACCTGGTGTGTACCTGCTGGAGTCACTCCAGCGGGGCTCAGACCTAC 700
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
Db 701 GAGATGTTCTCTGTTCTGCGGCCCTGCGCGGCCCTGCCAGGCCCGCCACAGAGCTCTCCACAC 760
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260

```
Db 761 GTCAGCACCTTGGTGCCTCCACAGTCGGCAGGTGCAGCAGCAGGGAACTGGAT 820
Qy 261 AlaSerGlyProAlaAspLeuValAspAlaPheLeuLeuLeuMetAlaGlnGlu 280
Db 821 GCTTGGGCCCCGACGTGACCTTGTGATGCCTTCTGCTGAAGATGGCAGAGGAA 880
Qy 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db 881 CAAAACCCAGGCACAGAAATTCACCAACAAGAACATGCTGATGACAGTCAATTTTGTG 940
Qy 301 PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMetLys 320
Db 941 TTTCTGGGAGGATGACGGTCAGCACCGGTGCGCTATACCTCTCTCTCTGATGAA 1000
Qy 321 TyrProHisValGlnLysTyrValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
Db 1001 TACCTCATGTCCTCAAAAGTGGTACGTGAGGAGCTGAATCGGAGCTGGGGCTGGCCAG 1060
Qy 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db 1061 GCACCAAGCCTAGGGACCGTACCGGCTCCCTTACACCGAGCGGTTCGATGAGGCG 1120
Qy 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrArg 380
Db 1121 CAGCGGCTGTGGCGCTGGTGCCTATGGGAATACCCGACACCTCATGCGAACCACCGC 1180
Qy 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db 1181 TTCGAGGTACACCTTGCCCGAGGACGAGGTCTTCCCTCTCTTGGCTCCATCTG 1240
Qy 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1241 CATGACCCCAACATCTTCAAGCACCCAGAGAGTTCACCCAGACCGTTTCTCGATGCA 1300
Qy 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1301 GATGGACGGTTCAGGAAGCATGAGGCTTCTGCTCTCTTCTAGGGAAGCGTGTCTG 1360
Qy 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPhePheThrIleLeuGlnAla 460
Db 1361 CTTGGAGAGGCTTGCAAAAGCGGAGCTTCTCTCTTCTTCCACCATCTTACAGGC 1420
Qy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1421 TTCTCTCTGAGAGCGCTGCCCGCGGACACCTTGAGCTCAAGCCACCGTCAGTGGC 1480
Qy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1481 CTTTTCAACATTCCTCCAGCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1540
Qy 501 ThrGlnThrArg 504
Db 1541 ACGCAGACCCAGA 1552

RESULT 9
ACA67025
ID ACA67025 standard; cDNA; 2623 BP.
XX
AC ACA67025;
XX
DT 23-JUN-2003 (first entry)
XX
DE cDNA encoding human PRO polypeptide #32.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein;
KW anti-PRO antibody; diagnostic assay; gene expression; diabetes;
KW bone disorder; cartilage disorder; rheumatoid arthritis; obesity;
KW sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;
KW hearing loss; coagulation disorder; stroke; heart attack; cardiac;
KW antidiabetic; anorectic; vulnery; antiarthritic; osteopathic;
KW antirheumatic; auditory; cerebroprotective; angiogenic; gene; ss.
XX
OS Homo sapiens.
```

```
XX US20030043111-Al.
PN 19-DEC-2001; 2001US-00028072.
XX 18-JUN-1997; 97US-0049911P.
XX 26-AUG-1997; 97US-0056974P.
XX 17-SEP-1997; 97US-0059113P.
XX 17-SEP-1997; 97US-0059115P.
XX 17-SEP-1997; 97US-0059117P.
XX 17-SEP-1997; 97US-0059122P.
XX 17-SEP-1997; 97US-0059184P.
XX 18-SEP-1997; 97US-0059263P.
XX 19-SEP-1997; 97US-0059352P.
XX 19-SEP-1997; 97US-0059588P.
XX 24-SEP-1997; 97US-0059836P.
XX 17-OCT-1997; 97US-0062250P.
XX 17-OCT-1997; 97US-0062285P.
XX 17-OCT-1997; 97US-0062287P.
XX 17-OCT-1997; 97US-0063755P.
XX 24-OCT-1997; 97US-0062814P.
XX 24-OCT-1997; 97US-0062816P.
XX 24-OCT-1997; 97US-0063045P.
XX 24-OCT-1997; 97US-0063082P.
XX 24-OCT-1997; 97US-0063127P.
XX 27-OCT-1997; 97US-0063327P.
XX 27-OCT-1997; 97US-0063329P.
XX 28-OCT-1997; 97US-0063550P.
XX 28-OCT-1997; 97US-0063561P.
XX 29-OCT-1997; 97US-0063704P.
XX 29-OCT-1997; 97US-0063733P.
XX 29-OCT-1997; 97US-0063735P.
XX 29-OCT-1997; 97US-0063738P.
XX 03-NOV-1997; 97US-0064488P.
XX 07-NOV-1997; 97US-0064809P.
XX 12-NOV-1997; 97US-0065186P.
XX 17-NOV-1997; 97US-0065846P.
XX 21-NOV-1997; 97US-0066364P.
XX 24-NOV-1997; 97US-0066453P.
XX 24-NOV-1997; 97US-0066511P.
XX 24-NOV-1997; 97US-0066770P.
XX 11-DEC-1997; 97US-0069212P.
XX 11-DEC-1997; 97US-0069278P.
XX 11-DEC-1997; 97US-0069334P.
XX 16-DEC-1997; 97US-0069694P.
XX 23-JAN-1998; 98US-0072320P.
XX 04-FEB-1998; 98US-0073612P.
XX 09-FEB-1998; 98US-0074086P.
XX 09-FEB-1998; 98US-0074092P.
XX 12-MAR-1998; 98US-0077791P.
XX 20-MAR-1998; 98US-0078910P.
XX 25-MAR-1998; 98US-0079294P.
XX 27-MAR-1998; 98US-0079663P.
XX 27-MAR-1998; 98US-0079728P.
XX 31-MAR-1998; 98US-0080165P.
XX 12-JUN-1998; 98WO-US012456.
XX 14-JUL-1998; 98WO-US014552.
XX 28-AUG-1998; 98WO-US017888.
XX 10-SEP-1998; 98WO-US018824.
XX 14-SEP-1998; 98WO-US019093.
XX 14-SEP-1998; 98WO-US019094.
XX 14-SEP-1998; 98WO-US019177.
XX 16-SEP-1998; 98WO-US019330.
XX 17-SEP-1998; 98WO-US019437.
XX 07-OCT-1998; 98WO-US021141.
XX 29-OCT-1998; 98WO-US022991.
XX 29-OCT-1998; 98WO-US022992.
XX 20-NOV-1998; 98WO-US024855.
XX 01-DEC-1998; 98WO-US025108.
XX 05-JAN-1999; 99WO-US000106.
XX 08-MAR-1999; 99WO-US005028.
```

Alignment Scores:	5.37e-231	Length:	2623
Pred. No.:		Matches:	504
Score:	2615.00	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%		


```
Db      1061 GCACCAAGCCTAGGGAGCGGTACCCGCTCCCTTACACCGAGCGGTTCTGATGAGGGC 1120
Qy      361 GlnArgLeuLeuAlaLeuValProMetGlyLeuProArgThrLeuMetArgThrArg 380
Db      1121 CAGCGGCTGCTGGCGCTGGTGGCCATGGGAATACCCGACCCCTCATGGGACCAACCCGC 1180
Qy      381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPhePheProLeuLeuGlySerIleLeu 400
Db      1181 TTCGAGGGTACACCTGCCCGCAGGGCAGGAGTCTTCCCCCTCCTTGGCTCCATCCTG 1240
Qy      401 HisAspProAsnIlePheLysHisProGluLysPheAsnProAspArgPheLeuAspAla 420
Db      1241 CATGACCCCAACATCTTCAAGCACCACAGAGATTCAACCCAGACCGTTTCTGGATGCA 1300
Qy      421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db      1301 GATGACCGTTTCAGAGACATGAGCGGTCTCTGCTTCTCTAGGAAGCGTGTCTGC 1360
Qy      441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPhePheThrThrIleLeuGlnAla 460
Db      1361 CTTGAGAGGCGCTGGCAAGAGCGGAGCTCTTCTTCTTTCACACCATCTCAAGCC 1420
Qy      461 PheSerLeuGluSerProCysProAspThrLeuSerLeuLysProThrValSerGly 480
Db      1421 TTCTCCCTGGAGAGCGCGTGGCCGCGGACACCTGAGCCTCAAGCCACCGTCAGTGC 1480
Qy      481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db      1481 CTYTTCACATCTCCCGAGCCTTCCAGTGCAGTCCGCTCCCACTGACCTTCACTCCACC 1540
Qy      501 ThrGlnThrArg 504
Db      1541 ACGCAGACCCAGA 1552

RESULT 10
ACA03634
ID ACA03634 standard; cDNA; 2623 BP.
XX
AC ACA03634;
XX
DT 23-MAY-2003 (first entry)
XX
DE cDNA encoding human PRO polypeptide #32.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein;
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;
KW differentiation; chondrocyte; tumour; genetic disorder; cytostatic; gene;
KW ss.
XX
OS Homo sapiens.
XX
PN US2003036180-A1.
XX
PD 20-FEB-2003.
XX
PF 09-MAY-2002; 2002US-00143114.
XX
PR 31-MAR-1997; 98WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US000376.
PR 18-FEB-2000; 2000WO-US000365.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
```


QY 441 LeuGlyGluGlyLeuAlaLysalaGluLeuPhePhePheThrThrileLeuGlnAla 460
Db 1361 CTTGAGAGGGCTCGCAAAAGCGAGCTCTTCTTCTTCCACCATCTCAAGCC 1420
QY 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1421 TTCCTCTGGAGAGCGCTGCGCGCGGACACCTTGAGCTCAAGCCACCGTCAGTGGC 1480
QY 481 LeuPheAenIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1481 CTTTTCACATTCCTCCAGCCTTCCAGCTGCAGTCCGTCCTCCACTGACCTTCACTCCACC 1540
QY 501 ThrGlnThrArg 504
Db 1541 ACGCAGACCAGA 1552
RESULT 11
ABX89172
ID ABX89172 standard; cDNA; 2623 BP.
XX
AC ABX89172;
XX
DT 13-MAY-2003 (first entry)
XX
DE DNA encoding novel secreted and transmembrane protein PRO1906.
XX
KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumour; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
KW retinal neurons cell survival; rod photoreceptor cell survival;
KW retinal disorder; retinitis pigmentosa; kidney disorder;
KW mammalian kidney mesangial cell proliferation; Berger disease;
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
KW chondrocyte redifferentiation; sports injury; arthritis; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003017563-A1.
XX
PD 23-JAN-2003.
XX
PF 07-MAY-2002; 2002US-00140808.
XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 16-SEP-1998; 98WO-US019177.
PR 17-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98WO-US019437.
PR 29-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 20-NOV-1998; 98WO-US023992.
PR 01-DEC-1998; 98WO-US024855.
PR 05-JAN-1999; 98WO-US025108.
PR 08-MAR-1999; 98WO-US000106.
PR 10-MAR-1999; 98WO-US005028.
PR 20-APR-1999; 98WO-US005190.
PR 14-MAY-1999; 98WO-US008615.
PR 02-JUN-1999; 98WO-US010733.
PR 01-SEP-1999; 98WO-US012252.
PR 08-SEP-1999; 98WO-US020111.
PR 13-SEP-1999; 98WO-US020594.
PR 15-SEP-1999; 98WO-US020944.
PR 15-SEP-1999; 98WO-US021090.
PR 05-OCT-1999; 98WO-US021547.
PR 29-NOV-1999; 98WO-US023089.
PR 98WO-US028214.
PR 30-NOV-1999; 98WO-US028313.
PR 30-NOV-1999; 98WO-US028409.
PR 01-DEC-1999; 98WO-US028301.
PR 01-DEC-1999; 98WO-US028634.
PR 02-DEC-1999; 98WO-US028551.
PR 02-DEC-1999; 98WO-US028584.
PR 02-DEC-1999; 98WO-US028585.
PR 16-DEC-1999; 98WO-US030095.
PR 20-DEC-1999; 98WO-US030911.
PR 22-DEC-1999; 98WO-US030999.
PR 22-DEC-1999; 98WO-US030720.
PR 30-DEC-1999; 98WO-US031243.
PR 30-DEC-1999; 98WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 10-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006319.
PR 20-MAR-2000; 2000WO-US006884.
PR 21-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US007532.
PR 17-MAY-2000; 2000WO-US008439.
PR 22-MAY-2000; 2000WO-US013705.
PR 30-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US00796498.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001WO-US0062706.
PR 14-MAR-2001; 2001WO-US008689.
PR 22-MAR-2001; 2001WO-US016744.
PR 05-APR-2001; 2001WO-US028366.
PR 10-MAY-2001; 2001WO-US0854208.
PR 18-MAY-2001; 2001WO-US0854280.
PR 25-MAY-2001; 2001WO-US0860216.
PR 25-MAY-2001; 2001WO-US0866028.
PR 25-MAY-2001; 2001WO-US0866034.
PR 01-JUN-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001WO-US0872035.
PR 05-JUN-2001; 2001WO-US017800.
PR 14-JUN-2001; 2001WO-US0874503.
PR 19-JUN-2001; 2001WO-US0882636.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001WO-US0887879.
PR 22-JUN-2001; 2001WO-US0201116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001WO-US008827.
PR 06-AUG-2001; 2001WO-US0924419.
PR 09-AUG-2001; 2001WO-US0927796.
PR 16-AUG-2001; 2001WO-US0931836.
PR 19-DEC-2001; 2001WO-US0028072.
XX
PA (GETH) GENENTECH INC.


```

Db      1361  CTTGGAGAGGCGCTGCAAAAGCGAGCTCTTCTCTTCAACCACTCCTCAAGCC 1420
Qv      461  PheSerLeuGluSerProCysProAspThrLeuSerLeuLeuSerProThrValSerGly 480
          |||||
Db      1421  TTTCTCCCTGGAGAGCGCTGCCCGCGGACACCTTGAGCGCTCAAGCCACCGTCAGTGCC 1480
          |||||
Qv      481  LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
          |||||
Db      1481  CTTTTCACATTCGCCAGCCTTCCAGCTGCCAGTCCGTCCTCCACCTGACCTTCACCTCCACC 1540
          |||||
Qv      501  ThrGlnThrArg 504
          |||||
Db      1541  ACGCAGACCAGA 1552

RESULT 12
ACD41826
ID      ACD41826 standard; cDNA; 2623 BP.
XX
AC      ACD41826;
XX
DT      05-SEP-2003 (first entry)
XX
DE      Human secreted/transmembrane protein (PRO) cDNA #32.
XX
KW      Human; ss; gene; PRO; secreted protein; transmembrane protein; tumour;
KW      cytostatic; gene therapy; tumour necrosis factor-alpha; TNF-alpha; blood;
KW      proteoglycan; cartilage; cytokine; peripheral blood mononuclear cell;
KW      PMBC; glucose uptake; FFA; skeletal muscle cell; adipocyte cell;
KW      chondrocyte cell proliferation; chondrocyte cell differentiation;
KW      pericyte cell; inner ear utricular supporting cell; T-lymphocyte cell;
KW      endothelial cell; A-peptide; factor VIIA.
XX
OS      Homo sapiens.
XX
PN      US2003036179 A1.
XX
PD      20-FEB-2003.
XX
PF      10-MAY-2002; 2002US-00142431.
XX
PR      31-MAR-1997; 99WO-US005230.
PR      12-JUN-1998; 98WO-US012456.
PR      14-JUL-1998; 98WO-US014552.
PR      28-AUG-1998; 98WO-US017888.
PR      10-SEP-1998; 98WO-US018824.
PR      14-SEP-1998; 98WO-US019093.
PR      14-SEP-1998; 98WO-US019094.
PR      16-SEP-1998; 98WO-US019177.
PR      17-SEP-1998; 98WO-US019330.
PR      07-OCT-1998; 98WO-US021141.
PR      29-OCT-1998; 98WO-US022991.
PR      29-OCT-1998; 98WO-US022992.
PR      01-DEC-1998; 98WO-US024855.
PR      05-JAN-1999; 99WO-US000106.
PR      08-MAR-1999; 99WO-US005028.
PR      10-MAR-1999; 99WO-US005190.
PR      20-APR-1999; 99WO-US008615.
PR      14-MAY-1999; 99WO-US010733.
PR      02-JUN-1999; 99WO-US012252.
PR      01-SEP-1999; 99WO-US020111.
PR      08-SEP-1999; 99WO-US020594.
PR      13-SEP-1999; 99WO-US020944.
PR      15-SEP-1999; 99WO-US021090.
PR      15-SEP-1999; 99WO-US021547.
PR      05-OCT-1999; 99WO-US023089.
PR      29-NOV-1999; 99WO-US028214.
PR      30-NOV-1999; 99WO-US028313.
PR      30-NOV-1999; 99WO-US028409.
PR      01-DEC-1999; 99WO-US028301.
PR      01-DEC-1999; 99WO-US028634.
PR      02-DEC-1999; 99WO-US028551.
PR      02-DEC-1999; 99WO-US028564.
PR      02-DEC-1999; 99WO-US028565.
PR      16-DEC-1999; 99WO-US030095.
PR      20-DEC-1999; 99WO-US030911.
PR      20-DEC-1999; 99WO-US030999.
PR      22-DEC-1999; 99WO-US030720.
PR      30-DEC-1999; 99WO-US031243.
PR      30-DEC-1999; 99WO-US031274.
PR      05-JAN-2000; 2000WO-US000219.
PR      06-JAN-2000; 2000WO-US000277.
PR      11-FEB-2000; 2000WO-US003565.
PR      18-FEB-2000; 2000WO-US004341.
PR      22-FEB-2000; 2000WO-US004414.
PR      24-FEB-2000; 2000WO-US004914.
PR      24-FEB-2000; 2000WO-US005004.
PR      01-MAR-2000; 2000WO-US005601.
PR      02-MAR-2000; 2000WO-US005746.
PR      02-MAR-2000; 2000WO-US005841.
PR      10-MAR-2000; 2000WO-US006319.
PR      15-MAR-2000; 2000WO-US006884.
PR      20-MAR-2000; 2000WO-US007377.
PR      31-MAR-2000; 2000WO-US007532.
PR      17-MAY-2000; 2000WO-US008439.
PR      22-MAY-2000; 2000WO-US013705.
PR      30-MAY-2000; 2000WO-US014042.
PR      02-JUN-2000; 2000WO-US014941.
PR      28-JUL-2000; 2000WO-US015264.
PR      11-AUG-2000; 2000WO-US020710.
PR      23-AUG-2000; 2000WO-US022031.
PR      24-AUG-2000; 2000WO-US023522.
PR      08-NOV-2000; 2000WO-US030952.
PR      10-NOV-2000; 2000WO-US030873.
PR      01-DEC-2000; 2000WO-US032678.
PR      20-DEC-2000; 2000US-0074259.
PR      28-DEC-2000; 2000WO-US034956.
PR      28-FEB-2001; 2001US-00796498.
PR      01-MAR-2001; 2001WO-US006520.
PR      09-MAR-2001; 2001WO-US006666.
PR      14-MAR-2001; 2001US-00802706.
PR      22-MAR-2001; 2001US-00808689.
PR      05-APR-2001; 2001US-00828366.
PR      10-MAY-2001; 2001US-00854208.
PR      18-MAY-2001; 2001US-00854280.
PR      25-MAY-2001; 2001US-00860216.
PR      25-MAY-2001; 2001US-00866026.
PR      25-MAY-2001; 2001US-00866034.
PR      25-MAY-2001; 2001WO-US017092.
PR      01-JUN-2001; 2001US-00872035.
PR      01-JUN-2001; 2001WO-US017800.
PR      05-JUN-2001; 2001US-00874503.
PR      14-JUN-2001; 2001US-00882636.
PR      19-JUN-2001; 2001US-00886342.
PR      20-JUN-2001; 2001WO-US019692.
PR      21-JUN-2001; 2001US-00887879.
PR      22-JUN-2001; 2001WO-US020116.
PR      29-JUN-2001; 2001WO-US021066.
PR      09-JUL-2001; 2001WO-US021735.
PR      18-JUL-2001; 2001US-00908827.
PR      06-AUG-2001; 2001US-00924419.
PR      09-AUG-2001; 2001US-00927796.
PR      16-AUG-2001; 2001US-00931836.
PR      19-DEC-2001; 2001US-00028072.
XX
XX      (GETH ) GENENTECH INC.
XX
PI      Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI      Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI      Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX

```


Db 1301 GATGAGCGTTGAGGAGCATGAGCGTTCCTGCGCCCTTCTCTAGGAGCGTGTCTGC 1360
 QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePheThrThrileLeuGlnAla 460
 Db 1361 CTTGAGAGGGCTTGGCAAGCGAGCTTCTTCTTCCACCACTCCTCAAGCC 1420
 QY 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480
 Db 1421 TTCTCCTGGAGAGCGCTGCGCGCGGACACCTTGAGCCTCAAGCCACCGTCAAGTGC 1480
 QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
 Db 1481 CTTTTCACATTTCCCGAGCCTTCCAGCTGCAAGTCCGTCCTCCACTGACCTTCACTCCACC 1540
 QY 501 ThrGlnThrArg 504
 Db 1541 ACGCAGACCAGA 1552

RESULT 13
 ACRA04055
 ID ACRA04055 standard; cDNA; 2623 BP.
 XX
 AC ACRA04055;
 XX
 DT 27-MAY-2003 (first entry)
 XX
 DE Human cDNA encoding a secreted/transmembrane protein, SEQ ID 63.
 XX
 KW Human; ss; gene; secreted protein; transmembrane protein; PRO;
 KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;
 KW infertility; birth defects; premature aging; AIDS; biosensor;
 KW acquired immunodeficiency syndrome; cancer; diabetic complication;
 KW bioreactor; tumour.
 XX
 OS Homo sapiens.
 XX
 PN US2003032155-A1.
 XX
 PD 13-FEB-2003.
 XX
 PF 03-MAY-2002; 2002US-00137865.
 XX
 PR 31-MAR-1997; 99WO-US005230.
 PR 12-JUN-1998; 98WO-US012456.
 PR 14-JUL-1998; 98WO-US014552.
 PR 28-AUG-1998; 98WO-US017888.
 PR 10-SEP-1998; 98WO-US018824.
 PR 14-SEP-1998; 98WO-US019093.
 PR 14-SEP-1998; 98WO-US019094.
 PR 16-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98WO-US019437.
 PR 07-OCT-1998; 98WO-US021141.
 PR 29-OCT-1998; 98WO-US022991.
 PR 29-OCT-1998; 98WO-US022992.
 PR 20-NOV-1998; 98WO-US024855.
 PR 01-DEC-1998; 98WO-US025108.
 PR 05-JAN-1999; 98WO-US000106.
 PR 08-MAR-1999; 98WO-US005028.
 PR 10-MAR-1999; 98WO-US005190.
 PR 20-APR-1999; 98WO-US008615.
 PR 14-MAY-1999; 98WO-US010733.
 PR 02-JUN-1999; 98WO-US012252.
 PR 01-SEP-1999; 98WO-US020111.
 PR 08-SEP-1999; 98WO-US020594.
 PR 13-SEP-1999; 98WO-US020944.
 PR 15-SEP-1999; 98WO-US021090.
 PR 15-SEP-1999; 98WO-US021547.
 PR 05-OCT-1999; 98WO-US023089.
 PR 29-NOV-1999; 98WO-US028214.
 PR 30-NOV-1999; 98WO-US028313.
 PR 30-NOV-1999; 98WO-US028409.
 PR 01-DEC-1999; 98WO-US028301.

PR 01-DEC-1999; 99WO-US028634.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028564.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 20-DEC-1999; 99WO-US030999.
 PR 22-DEC-1999; 99WO-US030720.
 PR 30-DEC-1999; 99WO-US031243.
 PR 05-JAN-2000; 99WO-US031274.
 PR 06-JAN-2000; 99WO-US000219.
 PR 06-JAN-2000; 99WO-US000277.
 PR 11-FEB-2000; 99WO-US000376.
 PR 18-FEB-2000; 99WO-US000365.
 PR 18-FEB-2000; 99WO-US004341.
 PR 22-FEB-2000; 99WO-US004342.
 PR 24-FEB-2000; 99WO-US004414.
 PR 24-FEB-2000; 99WO-US004914.
 PR 24-FEB-2000; 99WO-US005004.
 PR 01-MAR-2000; 99WO-US005601.
 PR 02-MAR-2000; 99WO-US005746.
 PR 10-MAR-2000; 99WO-US005841.
 PR 15-MAR-2000; 99WO-US006319.
 PR 20-MAR-2000; 99WO-US006884.
 PR 21-MAR-2000; 99WO-US007377.
 PR 30-MAR-2000; 99WO-US007532.
 PR 17-MAY-2000; 99WO-US008439.
 PR 22-MAY-2000; 99WO-US013705.
 PR 30-MAY-2000; 99WO-US014042.
 PR 02-JUN-2000; 99WO-US014941.
 PR 28-JUL-2000; 99WO-US015264.
 PR 11-AUG-2000; 99WO-US020710.
 PR 23-AUG-2000; 99WO-US023522.
 PR 24-AUG-2000; 99WO-US023328.
 PR 08-NOV-2000; 99WO-US030952.
 PR 10-NOV-2000; 99WO-US030956.
 PR 20-DEC-2000; 99WO-US032678.
 PR 01-DEC-2000; 99WO-US030873.
 PR 28-FEB-2001; 99WO-US00796498.
 PR 01-MAR-2001; 99WO-US006520.
 PR 03-MAR-2001; 99WO-US006666.
 PR 14-MAR-2001; 99WO-US0082706.
 PR 22-MAR-2001; 99WO-US008689.
 PR 05-APR-2001; 99WO-US016744.
 PR 10-MAY-2001; 99WO-US028366.
 PR 10-MAY-2001; 99WO-US0854208.
 PR 18-MAY-2001; 99WO-US0854280.
 PR 25-MAY-2001; 99WO-US0860216.
 PR 25-MAY-2001; 99WO-US0866028.
 PR 25-MAY-2001; 99WO-US0866034.
 PR 01-JUN-2001; 99WO-US017092.
 PR 01-JUN-2001; 99WO-US0872035.
 PR 05-JUN-2001; 99WO-US017800.
 PR 14-JUN-2001; 99WO-US0874503.
 PR 19-JUN-2001; 99WO-US0882636.
 PR 20-JUN-2001; 99WO-US0886342.
 PR 21-JUN-2001; 99WO-US019692.
 PR 22-JUN-2001; 99WO-US0887879.
 PR 29-JUN-2001; 99WO-US020116.
 PR 09-JUL-2001; 99WO-US021066.
 PR 18-JUL-2001; 99WO-US021735.
 PR 06-AUG-2001; 99WO-US0908827.
 PR 09-AUG-2001; 99WO-US0924419.
 PR 16-AUG-2001; 99WO-US0927796.
 PR 19-DEC-2001; 99WO-US0931836.
 PR 19-DEC-2001; 99WO-US0931836.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;

PI	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z;	Db	281	CTGGTTGGCAGGAGGCTGTGCGGAGGCCCTGGAGGTACGGCTGAGGAGTTTACGGCGC	340
XX	WPI; 2003-331925/31.	Qy	101	ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn	120
DR	P-PSDB; ABU66877.	Db	341	CGGGAAACCGTAGCGATGCTGGAAGGACTTTTGTATGGCCATGGGGTTTCTTCTCCAAC	400
XX	New secreted and transmembrane nucleic acids and polypeptides, designated	Qy	121	GlyGluArgTTPArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet	140
PT	as PRO, useful for treating inflammation, organ failure, atherosclerosis,	Db	401	GGGGAGCGGTGGAGGCAGCTGAGGAAGTTTACCATGCTTGTCTCTGGGACCTGGGCATG	460
PT	cardiac injury, infertility, birth defects, premature aging, AIDS, or	Qy	141	GlyLysArgGluGlyGluGluLeuIleGlnAlaArgCysLeuValGlnThrPhe	160
XX	cancer.	Db	461	GGGAACGAGAAGGCGCAGGAGCTGATCCAGGCGGAGGCCCGGTGCTGTGTGGAGCATTC	520
PS	Claim 2; Fig 63; 659pp; English.	Qy	161	GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn	180
CC	The invention relates to an isolated nucleic acid comprising, or which is	Db	521	CAGGGGACAGAAAGGACGCCATTCCATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAAC	580
CC	at least 80% identical to, or the full-length coding sequence of, any of	Qy	181	ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla	200
CC	the 275 nucleotide sequences, encoding the corresponding PRO polypeptide	Db	581	GTAGTCTGCTCCCTCTCTTGGCCCTCCCTCTCTATGAGGATAAGAGTTTCCAGGCC	640
CC	(one of 275 secreted or transmembrane proteins). The nucleic acid further	Qy	201	ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr	220
CC	comprises the full-length coding sequence of the DNA deposited under	Db	641	GTGGTCCGGGACGTGGTGGTACCTGTGGGAGTCAGCTCCAGGGGGGTTCAGACTAC	700
CC	American type Culture Collection (ATCC) accession number in a list given	Qy	221	GluMetPheSerTyrPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis	240
CC	in the specification. Also included are vectors and host cells for	Db	701	GAGATGTTCTCTGTTCTGTTCTGGGCCCTCCAGGCCCTCCACAGCAGCTCTCCACCAC	760
CC	producing PRO proteins, PRO fusion proteins, anti-PRO antibodies, PRO	Qy	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp	260
CC	extracellular domains and mature sequences, methods of detecting PRO	Db	761	GTCCAGCACCTTGGCTGCTTCCAGTCCGGCAGTGCAGCAGCAGCAGCAGCTCTCCAGG	820
CC	proteins, methods for stimulating the release of TNF-alpha (tumour	Qy	261	AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu	280
CC	necrosis, factor alpha) from human blood, (and the proliferation of	Db	821	GCTTCGGGCGCCGACGTCGCTGATGCTTCTCTGCTGAAGATGGCACAGGAGAA	880
CC	differentiation of chondrocyte cells, the proliferation of, or gene	Qy	281	GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu	300
CC	expression in pericyte cells, the release or proteoglycans from	Db	881	CAAAACCCAGGCACAGAAATCCCAACAGAACATCTGATGACAGTCAITTTATTGCTG	320
CC	cartilage, proliferation of inner ear articular supporting cells, the	Qy	301	PheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuLeuMetLys	340
CC	proliferation of T-lymphocyte cells, the release of a cytokine from	Db	941	TTTGCTGGGACGATGACGGTCAGCACCGTCCGCTATACCTCTCTCTCTCTCTCTCTCT	1000
CC	peripheral blood mononuclear cells (PBMC), or the proliferation of	Qy	321	TyrProHisValGlnLysTyrValArgGluLeuAsnArgGluLeuGlyAlaGlyGln	340
CC	endothelial cells), a method for modulating the uptake of glucose or free	Db	1001	TACCTCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGGAGGCTGGGGGCTGGCAG	1060
CC	fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the	Qy	341	AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla	360
CC	binding of A-peptide to factor VIIA, or the differentiation of adipocyte	Db	1061	GCACCAAGCCTAGGGGACCGTACCCGCTCTCTTACACCGACGCGGTCTCTGATGAGCGC	1120
CC	cells, a method for detecting the presence of any of the nucleotide sequences cited	Qy	361	GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg	380
CC	oligonucleotide probe derived from any of the nucleotide sequences cited	Db	1121	CAGCGGCTGCTGGCGCTGGTGGCATGGGAATACCCGCGACCTCTATCGCGAGCACCCCGC	1180
CC	above. The nucleic acids and polypeptides are useful for treating	Qy	381	PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu	400
CC	inflammatory diseases, organ failure, atherosclerosis, cardiac injury,	Db	1181	TTCCGAGGGTACACCTCTGCCCCCAGGGCAGGAGGTCTTCCCTCTCTCTCTCTCTCTCT	1240
CC	infertility, birth defects, premature aging, AIDS (acquired	Qy	401	HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla	420
CC	immunodeficiency syndrome), cancer, or diabetic complications. The	Db	1241	CATGACCCCAACATCTTCAAGCAGCCCAAGAGAGTTCACCCAGACCGCTTCTCTGATGCA	1300
CC	nucleic acids are useful as hybridisation probes, in chromosome and gene	Qy	421	AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys	440
CC	mapping, and in generating antisense RNA or DNA. The polypeptides are	Db	1301	GATGGACGGTTCAGAAAGCATGAGGCGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT	1360
CC	useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both	Qy	441	LeuGlyGluGlyLeuAlaLysAlaGluLeuPheLeuPheThrThrIleLeuGlnAla	460
CC	are useful in tissue typing. The present sequence encodes a PRO protein	Db	1361	CTTGAGAGGGCGCTGGCAAAAGCGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1420
CC	of the invention				
XX					
SQ	Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;				
Alignment Scores:					
Pred. No.:	5,37e-231	Length:	2623		
Score:	2615.00	Matches:	504		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	7	Gaps:	0		
US-10-669-693-2 (1-504) x ACA04055 (1-2623)					
Qy	1	MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuThr	20		
Db	41	ATGAGGCGACCGGACCTGGCGCTGCTGCTGGCGTGGCGTGGCTGCTGCTGCTGCTGCTG	100		
Qy	21	LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuPro	40		
Db	101	CTGGCGTGTGCGGAGCAGGCGCCAGGCGCCAGGCGCCAGGCGCCAGGCGCCAGGCGCC	160		
Qy	41	LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu	60		
Db	161	CTGCTGGAAACCTCTCTGACGTACGCGCCGGCGCGCTGTATTGAGGCGCTCATGGCGTG	220		
Qy	61	SerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTyrArgProValValVal	80		
Db	221	AGTAAGAAGTACGAGCGGTGTTCCATCTACCTGGGACCTCTGGCGCGCTGTGGTGTGTC	280		
Qy	81	LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlnAlaGluGluPheSerGly	100		


```

Oy 461 PheSerLeuGluSerProCysProAspThrLeuSerLeuValSerProThrValSergly 480
Db 1421 TTCTCCCTGGAGAGCCCTGCCCGCGGACACCCCTGAGCCTCAAGCCACCGTCAGTGGC 1480
Oy 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1481 CTTTTCACATTCCCGCCAGCCTTCAGCTGCAAGTCCGTCCTCCACTGACCTTCATCCACC 1540
Oy 501 ThrGlnThrArg 504
Db 1541 ACGCAGACCAGA 1552

RESULT 14
ID ADA45582 standard; cDNA; 2623 BP.
XX
AC ADA45582;
XX
DT 20-NOV-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1906 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; gene; ss;
KW Tumour necrosis factor alpha release; TNF-alpha release;
KW glucose uptake modulator; FFA uptake modulator;
KW cell proliferation stimulator; cell differentiation stimulator;
KW cell differentiation inhibitor; cytokine release stimulator; tumour;
KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; chromosome mapping; gene mapping;
KW gene therapy; chromosome identification; chromosome marker.
XX
OS Homo sapiens.
XX
PN US2003022328-A1.
XX
PD 30-JAN-2003.
XX
PF 16-APR-2002; 2002US-00123904.
XX
PR 31-MAR-1997; 99WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 01-DEC-1998; 98WO-US024855.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030999.
PR 30-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 11-FEB-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 01-MAR-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 10-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006319.
PR 20-MAR-2000; 2000WO-US006884.
PR 21-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US007532.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 28-FEB-2001; 2001US-00796498.
PR 01-MAR-2001; 2001WO-US006520.
PR 09-MAR-2001; 2001WO-US006666.
PR 14-MAR-2001; 2001US-00802706.
PR 22-MAR-2001; 2001US-00808689.
PR 05-APR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 18-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 03-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX

```

DR WPI; 2003-584997/55.
DR P-PSDB; ADA45583.
XX
XX Novel secreted and transmembrane polypeptide for modulating biological
PT activity of cell expressing the polypeptide, identifying agonists or
PT antagonists of polypeptide, and as molecular weight markers.
XX
XX
PS Claim 2; Fig 63; 659pp; English.
XX
XX The invention describes 305 nucleic acids encoding PRO (secreted and
CC transmembrane) polypeptides (I). (I) is useful for stimulating the
CC release of TNF-alpha from human blood, for modulating the uptake of
CC glucose or FFA by skeletal muscle cells or adipocyte cells, for
CC stimulating the proliferation or differentiation of chondrocyte cells,
CC for stimulating the proliferation of or gene expression in pericyte
CC cells, for stimulating the release of proteoglycans from cartilage, for
CC stimulating the proliferation of T-lymphocyte cells, for stimulating
CC the release of a cytokine from PBMC cells, for inhibiting the binding of
CC A-peptide to factor VIIa, for inhibiting the differentiation of adipocyte
CC cells, for stimulating proliferation of endothelial cells, for detecting
CC the presence of tumour in a mammal. The tumour is lung, colon, breast,
CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes
CC are useful for isolating genomic and cDNA nucleotide sequences or
CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful
CC in assays to identify other proteins or molecules involved in binding
CC interaction. A polynucleotide (II) encoding (I) is useful in chromosome
CC and gene mapping, in generation of antisense RNA and DNA, in the
CC preparation of PRO polypeptide, for generating transgenic animals or
CC knockout animals which in turn are useful in the development and
CC screening of therapeutically useful reagents, in gene therapy, for
CC chromosome identification, as chromosome marker, and for generating
CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.
CC detecting its expression in specific cells, tissues or serum, and for
CC affinity purification of PRO from recombinant cell culture or natural
CC sources. (I) and (II) are useful for tissue typing. This sequence encodes
CC a novel human secreted and transmembrane PRO polypeptide.
XX
SQ Sequence 2623 BP; 552 A; 865 C; 704 G; 502 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,37e-231 Length: 2623
Score: 2615.00 Matches: 504
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-669-693-2 (1-504) x ADA45582 (1-2623)

QY 1 MetGluAlaThrGlyThrTrpAlaLeuLeuLeuAlaLeuLeuLeuLeuLeuThr 20
Db 41 ATGGAGGCGACCGGACCTGGCGCTGCTGGCGCTGGCGCTGCTGGCGCTGACG 100
QY 21 LeuAlaLeuSerGlyThrArgAlaArgGlyHisLeuProGlyProThrProLeuPro 40
Db 101 CTGGCGCTGTCGGGACCGAGGCGCCGAGGCGACCTGCCCCCGGCGCCGCGCTACCA 160
QY 41 LeuLeuGlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeu 60
Db 161 CTGCTGGGAAACCTCTGCGAGCTACGCGCGGCGCGCTGTATTTCAGGGCTCATGGCGCTG 220
QY 61 SerLysIysTyrGlyProValPheThrIleTyrLeuGlyProThrArgProValVal 80
Db 221 AGTAAGAAGTACGAGCGGTGTTCACATCTACCTGGGACCCCTGGGCGCGCTGGGTGTC 280
QY 81 LeuValGlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGly 100
Db 281 CTGGTGGGAGGAGGCTGTGGGAGGCGCTGGGAGGTCAGGCTGAGGAGTTTCAGCGGC 340
QY 101 ArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSerAsn 120
Db 341 CGGGGAACCGTAGCGATGCTGGAAGGACTTTTGTATGGCCATGGGGTTTCTCTCCAAC 400

QY 121 GlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMet 140
Db 401 GGGAGAGCGGTGGAGGACGCTGAGGAGTTTACCAGTCTGCTCGGGGACCTGGGCGATG 460
QY 141 GlyLysArgGluGlyGluGluLeuGlnAlaGluAlaArgCysLeuValGluThrPhe 160
Db 461 GGGAGCGGAGAGGCGAGGAGCTGATCAGCGGAGGCGCGGTCTCTGTGTGAGACATTC 520
QY 161 GlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSerAsn 180
Db 521 CAGGGGACAGAGGAGCGCCCATTCGATCCTCCCTGCTGGCGGACGACCTCCAC 580
QY 181 ValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAla 200
Db 581 GTAGTCTGCTCCCTCTCTTGGCCCTCGCTTCTCTATGAGATAAGGAGTTCAGGCGC 640
QY 201 ValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlyGlnThrTyr 220
Db 641 GTGGTCCGGGACGCTGGTGGTACCTGCTGGGAGTCAGCTCCAGGGGGGTCCAGACTAC 700
QY 221 GluMetPheSerTrpPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHis 240
Db 701 GAGATGTTCTCTGCTGGTCTCGGGCCCTGCGGAGGCGCCCAAGCAGCTCTCTCCACC 760
QY 241 ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAsp 260
Db 761 GTCAGCAGCTTGGCTGGCTTCACAGTCCGGCAGGTGCAGCAGCACCAGGGGAACTGGAT 820
QY 261 AlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlu 280
Db 821 GCTTCGGGCGCGCACGTCGACCTTGTGATGCTCTCTGCTGAAGATGCGACAGGAGAA 880
QY 281 GlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeu 300
Db 881 CAAACCCAGGACAGAAATTCACCAACAAGACATGCTGATGACAGTCAATTTATTGCTG 940
QY 301 PheAlaGlyThrMetThrValSerThrValGlyTyrThrLeuLeuLeuLeuMetLys 320
Db 941 TTTGCTGGGACGATGACGCTCAGCACCAGGTCGGCTATACCTCTCTCTGATGATAA 1000
QY 321 TyrProHisValGlnLysTrpValArgGluGluLeuAsnArgGluLeuGlyAlaGlyGln 340
Db 1001 TACCCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGGAGCTGGGGGTGGCCAG 1060
QY 341 AlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGluAla 360
Db 1061 GCACCAAGCTAGGGGACCGTACCCGCTCCCTTACACGACGCGGTTCTGCAATGAGCG 1120
QY 361 GlnArgLeuLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThrArg 380
Db 1121 CAGCGGCTGCTGGCGTGGTGGCAATACCCCGCACCTCATCGGACACACCGCG 1180
QY 381 PheArgGlyTyrThrLeuProGlnGlyThrGluValPheProLeuLeuGlySerIleLeu 400
Db 1181 TCCGAGGGTACACCTGCCCCCAGGAGGACGAGGTCTTCCCTCTCTGCTCCATCTCTG 1240
QY 401 HisAspProAsnIlePheLysHisProGluGluPheAsnProAspArgPheLeuAspAla 420
Db 1241 CATGACCCCAACATCTCAAGCACCCAGAGAGTTCAACCCAGACCGTTTCTTGGATGCA 1300
QY 421 AspGlyArgPheArgLysHisGluAlaPheLeuProPheSerLeuGlyLysArgValCys 440
Db 1301 GATGGACGGTTTCAGGAAGCATGAGGCGTCTCTCCCTTCTCTTAGGGAAGCGTGTCTG 1360
QY 441 LeuGlyGluGlyLeuAlaLysAlaGluLeuPhePhePheThrThrIleLeuGlnAla 460
Db 1361 CTTGGAGAGGCGCTGGCAAAAGCGAGCTCTTCTCTTCTTCCACCACTCTTCTTCAAGCC 1420
QY 461 PheSerLeuGluSerProCysProProAspThrLeuSerLeuLysProThrValSerGly 480
Db 1421 TTTCTCTGGAGAGCCCGTGGCGCGGACACCTTGAGCCTCAAGCCACCGTCAGTGGC 1480

OY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
Db 1481 CTTTCAACATTCCTCCAGCTTCAGCTGCAAGTCCGTCCTCCACTGACCTTCCTCCACC 1540
OY 501 ThrGlnThrArg 504
Db 1541 ACGCAGACCAGA 1552
RESULT 15
ADA76013
ID ADA76013 standard; cDNA; 2623 BP.
XX
AC ADA76013;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human PRO polynucleotide #32.
XX
KW Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
KW tumour necrosis factor-alpha; TNF-alpha; chondrocyte cell; tumour;
KW cancer; adrenal; lung; colon; breast; prostate; rectum; kidney; cervix;
KW liver; microvascular endothelial cell; glucose; FFA;
KW skeletal muscle cell; adipocyte cell; pericyte cell;
KW inner ear utricular supporting cell; T-lymphocyte cell;
KW endothelial cell tube formation; bone disorder; cartilage disorder;
KW sports injury; proteoglycan; articular cartilage defect; osteoarthritis;
KW rheumatoid arthritis; haemoglobin-associated disorder thalassaemia;
KW immune system cell infiltration.
XX
OS Homo sapiens.
XX
PN US2003073212-A1.
XX
PD 17-APR-2003.
XX
PF 16-APR-2002; 2002US-00123903.
XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012452.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 01-DEC-1998; 98WO-US024855.
PR 05-JAN-1999; 98WO-US025108.
PR 08-MAR-1999; 98WO-US000106.
PR 10-MAR-1999; 98WO-US005028.
PR 20-APR-1999; 98WO-US005190.
PR 14-MAY-1999; 98WO-US008615.
PR 02-JUN-1999; 98WO-US010733.
PR 01-SEP-1999; 98WO-US021252.
PR 08-SEP-1999; 98WO-US020111.
PR 13-SEP-1999; 98WO-US020594.
PR 15-SEP-1999; 98WO-US020944.
PR 15-SEP-1999; 98WO-US021090.
PR 05-OCT-1999; 98WO-US021547.
PR 29-NOV-1999; 98WO-US023089.
PR 30-NOV-1999; 98WO-US028214.
PR 01-DEC-1999; 98WO-US028313.
PR 01-DEC-1999; 98WO-US028409.
PR 02-DEC-1999; 98WO-US028301.
PR 02-DEC-1999; 98WO-US028634.
PR 02-DEC-1999; 98WO-US028551.
PR 02-DEC-1999; 98WO-US028564.
PR 16-DEC-1999; 98WO-US028565.
PR 16-DEC-1999; 98WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 05-JAN-2000; 99WO-US031274.
PR 06-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 11-FEB-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014842.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001US-00866034.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-687639/65.
DR P-PSDB; ADA76014.
XX

Db 1421 TTCTCCTGGAGAGCCCGTSCCGCGGACACCTGAGCCTCAAGCCACCGTCAGTGGC 1480
 QY 481 LeuPheAsnIleProProAlaPheGlnLeuGlnValArgProThrAspLeuHisSerThr 500
 Db 1481 CTTTCAACATTCCCGCCAGCCTTCAGCTGCAAGTCCGTCCTGACCTTCACTCCACC 1540
 QY 501 ThrGlnThrArg 504
 Db 1541 ACGCAGACCAGA 1552

Search completed: September 18, 2004, 01:51:19
 Job time : 677 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 18, 2004, 00:15:11 ; Search time 3778 Seconds
(without alignments)
3983.732 Million cell updates/sec

Title: US-10-669-693-2
Perfect score: 2615
Sequence: 1 MEATGTWALLALALLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-Q=/cgm2_1/USPTO.spool/US10669693/runat_15092004_101310_5862/app_query.fasta_1.647
-DB=EST -QPMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10669693 @CGN 1 1 3437 @runat_15092004_101310_5862 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estmu:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gsl1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2060.5	78.8	2292	11	AK087069 Mus muscu
2	2060.5	78.8	2621	11	AK004699 Mus muscu
3	1529.5	58.5	900	13	BUS52396 AGENCOURT
4	1465.5	56.0	924	13	BUS49902 AGENCOURT
5	1447.5	55.3	903	13	BQ900302 AGENCOURT
6	1410.5	53.9	939	13	BQ958112 AGENCOURT
7	1376.5	52.6	2742	11	AK054324 Mus muscu
8	1362.5	52.1	857	12	BM047426 603628712
9	1321.5	50.5	877	13	BUS43186 AGENCOURT
10	1299.5	49.7	1876	11	AK028103 Mus muscu
11	1222.5	46.7	838	13	BQ957100 AGENCOURT
12	1146.5	43.8	1944	11	BC022141 Mus muscu
13	1138.5	43.5	969	13	EX422464 BX422464
14	1083.5	41.4	2870	11	AK050051 Mus muscu
15	1069.5	40.9	1870	11	AK008688 Mus muscu
16	1056.5	40.4	1825	11	AK005017 Mus muscu
17	1039.5	39.8	1567	14	CD013962 Mus muscu
18	1038.5	39.7	1990	11	AK008580 Mus muscu
19	1022.5	39.1	723	12	BI906340 603063307
20	1022.5	39.1	2022	14	CD014119 90127964
21	1018.5	38.9	2093	14	CD013967 Mus muscu
22	1007.5	38.5	1452	14	CD013965 Mus muscu
23	1005.5	38.4	2131	14	CD013968 Mus muscu
24	1001.5	38.3	2132	14	CD013969 90127948
25	998.5	38.2	1559	11	AK050436 Mus muscu
26	993.5	38.0	1482	29	AY407531 Mus muscu
27	993.5	38.0	1731	11	BC042693 Mus muscu
28	992.5	37.9	599	12	BM786289 K-EST0064
29	982.5	37.6	1482	29	AY407529 Homo sapi
30	972.5	37.2	1703	11	AK050082 Mus muscu
31	954.5	36.5	1339	14	CD013963 Mus muscu
32	953.5	36.4	1323	14	CD013966 Mus muscu
33	943.5	36.1	1299	14	CD013964 Mus muscu
34	936.5	35.8	1760	14	CD013970 Mus muscu
35	933.5	35.7	571	10	BE148597 MR0-HT024
36	928.5	35.5	648	12	BI016305 MR4-ET013
37	918.5	35.1	674	12	BI836998 603084539
38	911.5	34.9	669	13	BQ257072 NISC_K009
39	905.5	34.6	1773	11	AK008512 Mus muscu
40	903.5	34.5	538	12	BM795225 K-EST0076
41	893.5	34.1	548	13	BM646408 DXFZP781K
42	887.5	33.9	637	14	CF170039 B0821G03-
43	887.5	33.9	553	12	BM751056 K-EST0026
44	870.5	33.3	626	14	CF170325 B0825H09-
45	867.5	33.2	804	12	BM044172 603621408

ALIGNMENTS

RESULT 1
AK087069

LOCUS
DEFINITION

AK087069 2292 bp mRNA linear HTC 20-SEP-2003
Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched
library, clone:E030025C01 product:similar to CYTOCHROME P450 2S1
[Homo sapiens], full insert sequence.

ACCESSION

AK087069.1 GI:26352314
HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

QY	221	GluMetPheSerTrpPheLeuArgProLeuProGlyProHisGlnLeuLeuHisHis	240
Db	675	GAGATGTTCTCTGCTTCCTGCGGNCCTGCCAGGCCCCCAACAGCAGCTCTCTCCACCAC	734
QY	241	ValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGln-GlyAsnLeuAs	260
Db	735	GTCCAGCACCTTGGCTGCCTTCCACAGTCCGGCAGGTGCAGCAGCACCAAGGGAACCTGGGA	794
QY	260	pAlaSerGlyProAlaArgAspLeuValAspAlaPhe-LeuLeuLysMetAlaGlnGluG	280
Db	795	TGCTTTCGGGCCCGCACGTGAACGTGTGATGCCTTCCCTGCTGAAAGGCGCACCGAAG	854
QY	280	lueGlnAsnProGlyThrGlu-PheThrAsnLysAsnMetLeuMetThrValIleTyrLeu	299
Db	855	GACAAACCCAGGCACAGAAATTCACCAACAGAACATGCTGTGATGACATCATT-----	908
QY	300	LeuPheAlaGly 303	
Db	909	TTATTGTGCTGGG 920	
RESULT 5			
BQ900302		903 bp mRNA linear EST 16-AUG-2002	
LOCUS		AGENCOURT_8675726 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380105	
DEFINITION		5', mRNA sequence.	
ACCESSION		BQ900302	
VERSION		BQ900302.1 GI:229292316	
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		1 (bases 1 to 903)	
AUTHORS		NIH-MGC http://imgc.nci.nih.gov/.	
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cga@bbs-rcmail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov Plate: LLCM2568 row: c column: 18 High quality sequence stop: 690.	
FEATURES		Location/Qualifiers	
source		1..903	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6380105"	
		/tissue_type="carcinoma, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_40"	
		/notes="Organ: prostate; Vector: pOTB7; Site1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN			
Alignment Scores:			
Pred. No.:		8.12e-141	Length: 903
Score:		1447.00	Matches: 290
Percent Similarity:		96.6%	Conservative: 2
Best Local Similarity:		96.03%	Mismatches: 9
Query Match:		55.33%	Indels: 3
DB:		13	Gaps: 0

US-10-669-693-2 (1-504) x BQ900302 (1-903)	
QY	115 GlyValPhePheSerAsnGlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAla 134
Db	2 GGGGTTTCTCTCCAC -GGGAGCGGTGGAGCAGCTGAGAA-TTTACCATGCTTGCT 59
QY	135 LeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArg 154
Db	60 CTGCGGACCTGGGCATGGGAAGGAGAGCGAGAGCTGATCCAGCGGAGGCCCGG 119
QY	155 CysLeuValGluThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeu 174
Db	120 TGTCTGGTGGAGACATTCCAGGGGACAGAGAGCGCCATTCGATCCCTCCCTGCTGCTG 179
QY	175 AlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGlu 194
Db	180 GCCCAGGCCACCTCCAAAGTAGTCTGCTCCCTCTCTTGGCTCCGCTTCTCTATGAG 239
QY	195 AspLysGluPheGlnAlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSer 214
Db	240 GATAAGGAGTTCCAGGCGTGTCTCGGCGAGCTGTGTACCTCTGCTGGAGTCAGCTCC 299
QY	215 GlnGlyGlyGlnThrTyrGluMetPheSerTrpPheLeuArgProLeuProGlyProHis 234
Db	300 CAGGGGGTTCAGACTACGAGATGTTCTCTGGTTCCTGCGGCCCTGCCAGGCCCCAC 359
QY	235 LysGlnLeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGlnGln 254
Db	360 AAGCAGCTCTCCACCGCTCAGCACCTTGGCTGCTTCCTTCCAGTCCGCGAGTGCGAG 419
QY	255 HisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAlaPheLeuLeu 274
Db	420 CACCAGGGAACTGGATGCTTCGGGCCCGCGACGTGACCTTGTGATGCTTCTGCTG 479
QY	275 LysMetAlaGlnGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMet 294
Db	480 AAGATGCGCAGGAGGAAACAAACCCAGGACAGAAATTCACCAACAGACATGCTGATG 539
QY	295 ThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyrThr 314
Db	540 ACAGTCATTATTGCTGTTGCTGGAGAGATGACGGTCAGCACCGGTGGGTATACCC 599
QY	315 LeuLeuLeuMetLysTyrProHisValGlnLysTrpValArgGluGluLeuAsnArg 334
Db	600 CTCCTGCTCTGATGAATACCCCTCATGTCCAAAGTGGGTACGTGAGGAGCTGAATCGG 659
QY	335 GluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThrArgLeuProTyrThrAsp 354
Db	660 GAGCTGGGGCTGGCCAGGACCAAGCCTAGGGGACCGTACCGGCTCCCTTACACCGAC 719
QY	355 AlaValLeuHisGluAlaGlnArgLeuAlaLeuValProMetGlyIleProArgThr 374
Db	720 GCGGTTCTGATGAGCGCAGCAGGCTGCTGGCGCTGTGTCCTATGGAAATACCCCGACC 779
QY	375 LeuMetArgThrThrArgPheArgGlyTyrThrLeuProGlnGlyThrGluValPhe-Pr 394
Db	780 CTATCGGACCAACCCGCTTCCGAGGTAACACCTGCCCCCAGGACCGGAGGCTTCCCC 839
QY	394 oLeuLeuGlySerIleLeuHisAspProAsnIlePheLysHisProGluGluPheAsnPr 414
Db	840 CCTCTCTGGCTCATCTGATGACCCCACTCTTCCAGCACCCAGAGAGTTCACACC 899
QY	414 oAsp 415
Db	900 AGAC 903
RESULT 6	
BQ958112	
LOCUS	
DEFINITION AGENCOURT 10034258 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481143	
5', mRNA sequence.	
ACCESSION BQ958112	
US-10-669-693-2 (1-504) x BQ958112 (1-939)	
QY	115 GlyValPhePheSerAsnGlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAla 134
Db	2 GGGGTTTCTCTCCAC -GGGAGCGGTGGAGCAGCTGAGAA-TTTACCATGCTTGCT 60
QY	135 LeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeuLeuGlnAlaGluAlaArg 154
Db	61 CTCGCGGACCTGGGCATGGGAGCGAGAGCGAGAGCTGATCCAGCGGAGGCCCGG 120
QY	155 CysLeuValGluThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeu 174
Db	121 TGTCTGGTGGAGACATTCCAGGGGACAGAGAGCGCCATTCGATCCCTCCCTGCTGCTG 180
QY	175 AlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGlu 194
Db	181 GCCCAGGCCACCTCCAAAGTAGTCTGCTCCCTCTCTTGGCTCCGCTTCTCTATGAG 240
QY	195 AspLysGluPheGlnAlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSer 214
Db	241 GATAAGGAGTTCCAGGCCGCTGCTCGGCGAGCTGTGTGTTACCTCTGCTGGAGTCAGCTCC 300
QY	215 GlnGlyGlyGlnThrTyrGluMetPheSerTrpPheLeuArgProLeuProGlyProHis 234
Db	301 CAGGGGGGTTCAGACCTACGAGATGTTCTCTGCTGCTTCTCTGCGGCCCTGCCAGGCCCCAC 360
VERSION BQ958112.1 GI:22373590	
EST.	
SOURCE Homo sapiens (human)	
ORGANISM Homo sapiens	
REFERENCE 1 (bases 1 to 939)	
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.	
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL Unpublished (1999)	
COMMENT Contact: Robert Strausberg, Ph.D.	
Email: cgapbs@mail.nih.gov	
Tissue Procurement: DCTD/DTF	
cDNA Library Preparation: Rubin Laboratory	
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
Cloned by: Agencourt Bioscience Corporation	
Clone distribution: MGC clone distribution information can be	
found through the I.M.A.G.E. Consortium/LLNL at:	
http://image.llnl.gov	
Plate: LNCM2662 row: a column: 16	
High quality sequence stop: 636.	
FEATURES	
Location/Qualifiers	
1..939	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clones="IMAGE:6481143"	
/tissue_type="carcinoma, cell line"	
/lab_host="DH10B (phage-resistant)"	
/clone_lib="NIH_MGC_40"	
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;	
Site 2: EcoRI; cDNA made by oligo-dT priming.	
Directionally cloned into EcoRI/XhoI sites using the	
following 5' adaptor: GGCAGAG(G). Library constructed by	
Ling Hong in the laboratory of Gerald M. Rubin (University	
of California, Berkeley) using ZAP-cDNA synthesis kit	
(Stratagene) and Superscript II RT (Life Technologies).	
Note: this is a NIH_MGC Library."	
ORIGIN	
Alignment Scores:	
Pred. No.:	5,87e-137
Score:	1410.50
Percent Similarity:	95.42%
Best Local Similarity:	94.77%
Query Match:	53.94%
DB:	13
Gaps:	1

```
Qy 235 LysGlnLeuLeuHisHisValSerThrLeuAlaLalaPheThrValArgGlnValGlnGln 254
Db 361 AAGCAGCTCTCCACACGTCAGCACCTTGGCTGCTTTCACAGTCCGCGAGTGCAGCAG 420
Qy 255 HisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeu 274
Db 421 CACAGGGAACCTGGATGCTTCGGGGCCCGCAGCTGACCTTGTGTGATGCTTCTCTGCTG 480
Qy 275 LysMetAlaGlnGlnGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMet 294
Db 481 AAGATGGCAGAGGAGAACAAACCCAGGCACAGATTCCACCAACAGACATGCTGATG 540
Qy 295 ThrValIleTyLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyThr 314
Db 541 ACAGTCATTTATTTGCTGTTTCTGGGACGATGACGGTCAGCACCGTCCGGCTATACC 600
Qy 315 LeuLeuLeuLeuMetLysTyProHisValGlnLysTyrValArgGluLeuLeuAsnArg 334
Db 601 CTCCTGCTCTGATGAATACCTCATGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGG 660
Qy 335 GluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThrArgLeuProTyrThrAsp 354
Db 661 GAGCTGGGGGCTGGCCAGCAGCAACAGCTAGGGGACCGTACCGGCTCCCTTACACCGAC 720
Qy 355 AlaValLeuHisGlnAlaGlnArgLeuLeuAlaLeuValProMetGlyIleProArgThr 374
Db 721 GCGGNTCTGCATGAGCGGCANCGCTGCTGGCGCTGGTCCCATGGGAATACCCCGACCC 780
Qy 375 LeuMetArgThrThrArg-PheArgGlyTyThrLeuProGlnGly-ThrGlu---ValP 393
Db 781 CTCATCGGGACACCCGCTTTTCGAGGGTACACCTTGGCCCAAGNACCGNAGGTCCTT 840
Qy 393 heProLeuLeuGlySer-IleLeuHis-AspProAsn-IlePheLysHisProGluGluP 412
Db 841 CCCCCNCTTGGCTCCACTCTCGATGGACCCCAACAATCTTCAAGCACCCAGAAAGT 900
Qy 412 heAsnPro 414
Db 901 TTCAAGCC 908

RESULT 7
AK054324
LOCUS
DEFINITION
AK054324 2742 bp mRNA linear HTC 20-SEP-2003
Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
full-length enriched library, clone:E330015C15 product:similar to
CYTOCHROME P450 2S1 [Homo sapiens], full insert sequence.
ACCESSION AK054324
VERSION AK054324.1 GI:26096313
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED
REFERENCE
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
```

```
TITLE
JOURNAL MEDLINE
PUBMED
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2742)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,W.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source
1. .2742
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:E330015C15"
/db_xref="MGI:2426416"
/db_xref="taxon:10090"
/clone="E330015C15"
/sex="female"
/tissue_type="ovary"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days pregnant adult"
misc_feature
1. .2742
/note="similar to CYTOCHROME P450 2S1 [Homo sapiens]
(SPT9B266, evidence: FASTY, 78.4%ID, 99%length,
match=1500)"
ORIGIN
Alignment Scores:
Pred. No.: 1.24e-132 Length: 2742
Score: 1376.50 Matches: 268
Percent Similarity: 82.26% Conservativeness: 38
Best Local Similarity: 72.04% Mismatches: 43
Query Match: 52.64% Indels: 23
```

```

DB: 11 Gaps: 4
US-10-669-693-2 (1-504) x AK054324 (1-2742)

QY 60 LeuSerLysLysTyrGlyProValPheThrIleTyrLeuGlyProTyrArgProValVal 79
DB 33 CTAAGTAAGAAGTATGGGCTGTGTTCACGGTATACCTGGGCCCTTGGCGCGCGTGGTG 92
QY 80 ValLeuValGlyGlnAlaValArgGluAlaLeuGlyGlyGlnAlaGluPheSer 99
DB 93 GTCCTGGTGGACATGCTGTAGAGAGAGCTTGGGAGGTGAGCTCAGGAATTCAGC 152
QY 100 GlyArgGlyThrValAlaMetLeuGluGlyThrPheAspGlyHisGlyValPhePheSer 119
DB 153 GGGCGTGAACATTCGCAACGCTGACAGAGCTTTGATGGTCACGGAGTTTCTTGGC 212
QY 120 AsnGlyGluArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGly 139
DB 213 AATGGGAGCGGTGAAACAGCTGAGGAAATTCACCCCTGCTCGCTACGGGACCTGGGC 272
QY 140 MetGlyLysArgGlyGluGluLeuIleGlnAlaGluAlaArgCysLeuValGluThr 159
DB 273 ATGGCAGCCAGAGGCGAGAGCTGATCCAGCGGAGGTGCAGAGTCTGGTGAGGCT 332
QY 160 PheGlnGlyThrGluGlyArgProPheAspProSerLeuLeuAlaGlnAlaThrSer 179
DB 333 TTCAGAGAGCAGAGGAGGCTTCATTCACCTTCATGCTGCTGGCCCGAGCCACTCT 392
QY 180 AsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGln 199
DB 393 AATGTGCTGTCTCCCTGTCTTGGCATCGCTTGGCCCTATGACGATAAAGAGTTCCAG 452
QY 200 AlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSerSerGlnGlyGlnThr 219
DB 453 GCTGTGATCCAGGAGCAGAGTGTACTTGTAGGATCATCCTCTCCATGGGCGAGGCC 512
QY 220 TyrGluMetPheSerTyrPheLeuArgProLeuProGlyProHisLysGlnLeuHis 239
DB 513 TACAGATGTCTCTGGCTACTGCAGCCCTGCAGGCCCCACACACAGCTCCAGCAC 572
QY 240 HisValSerThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeu 259
DB 573 CACTGGGCGCCGTGGCTGCTTCCATTCACGAGGTACAGAAACACACAGGAGCGCTTC 632
QY 260 AspAlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGlu 279
DB 633 CAAACCTCAGTCTCGACGTGATGCTGTGACCCCTTCTGCTAAAGATGGCAGGAG 692
QY 280 GluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeu 299
DB 693 AAACAGACCCAGGTACAGAAATTCACCGAAGAACTTGTGTATGACGGTCAATACCTG 752
QY 300 LeuPheAlaGlyThrMetThrValSerThrThrValGlyTyrThrLeuLeuLeuMet 319
DB 753 CTGTTTGTGGGACCATGACCATCGGTGCGCCATCGCTATGCTGCTCTCTCTGCTG 812
QY 320 LysTyrProHisValGlnLysTyrValArgGluGluLeuAsnArgGluLeuGlyAlaGly 339
DB 813 AGATACCTCAAGTCCAGAGCGCGTCCGGAGAGAGCTCATACAGAGCTGGGTCTGGC 872
QY 340 GlnAlaProSerLeuGlyAspArgThrArgLeuProTyrThrAspAlaValLeuHisGlu 359
DB 873 AGGGCTCCAAGTCTACGCGATCGAGTTCGCTCCCTTACACGATGCGCTTTTACAGG 932
QY 360 AlaGlnArgLeuAlaLeuValProMetGlyIleProArgThrLeuMetArgThrThr 379
DB 933 GCACAGCGGCTCTGGCATGTGTACCATGAGGATGGGATGCCCATCCATCAGAGGACCACT 992
QY 380 ArgPheArgGlyTyrThrLeuProGln----- 388
DB 993 TGTCTCCGAGGTACTACTCTGCCAGGTAGGTAGGTATGGCAGCCAGAGCAATATCTC 1052
QY 389 -----GlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePhe 406

```

```

Db 1053 TGCCCTGGGACTTGGCTT-----TTGATATCCCTGTCTCTGGAGTACT----- 1097
QY 407 LysHisProGlu-----GluPheAsnPro 414
Db 1098 ---CACCCCAACCACTCCACCTTTTCTCTCC 1130

RESULT 8
BM047426 857 bp mRNA linear EST 07-NOV-2001
603628712F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5457075 5',
mRNA sequence.
ACCESSION BM047426
VERSION BM047426.1 GI:16776693
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 857)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1956 row: h column: 04
High quality sequence stop: 817.
FEATURES
Location/Qualifiers
1..857
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5457075"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 6,26e-132 Length: 857
Score: 1362.00 Matches: 277
Percent Similarity: 97.54% Conservative: 0
Best Local Similarity: 97.54% Mismatches: 7
Query Match: 52.08% Indels: 3
DB: 12 Gaps: 0
US-10-669-693-2 (1-504) x BM047426 (1-857)

QY 114 HisGlyValPhePheSerAsnGlyGluArgTyrArgGlnLeuArgLysPheThrMetLeu 133
Db 3 CATGGGTTTCTTCTCCAACGGGAGCGGTGGAGCAGCTGAGGAAGTTTACATGCTT 62
QY 134 AlaLeuArgAspLeuGlyMetGlyLysArgGluGlyGluGluLeuIleGlnAlaGluAla 153
Db 63 GCTCTGGGAGCTTGGCATGGGAGCGAGAGCGAGAGCTGATCCAGCGGAGGCC 122
QY 154 ArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeu 173

```


found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM2771 row: 1 column: 10

High quality sequence stop: 703.

FEATURES

source

1. 877

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6575194"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. NO.: 1.31e-127 Length: 877

Score: 1321.00 Matches: 271

Percent Similarity: 94.46% Conservative: 2

Best Local Similarity: 93.77% Mismatches: 10

Query Match: 50.52% Indels: 7

DB: 13 Gaps: 2

US-10-669-693-2 (1-504) x BU543186 (1-877)

QY 115 GlyValPhePheSerAsnGlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAla 134
 DB 2 GGGGTTTCTTCTCCAC-GGGAGCGGTGGAGCAGCTGAGGAAGTTTACCATGCTGCT 60
 QY 135 LeuArgAspLeuGlyMetGlyLysArgGluGluGluGluGluGluGluGluGluGlu 154
 DB 61 CTGCGGACCTGGGATCGGAGCGAGAGCGAGAGCTGATCCAGGCGAGGCGCCCG 120
 QY 155 CysLeuValGluThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeu 174
 DB 121 TGTCTGTGTGAGACATTCAGGGGACAGAGGACGCCCATTCGATCCCTCCCTGCTG 180
 QY 175 AlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGlu 194
 DB 181 GCCCAGGCCACCTCCCAACGTAGTCTGCTCCCTCTCTTTGGCCTCCGCTTCTCTATGAG 240
 QY 195 AspLysGluPheGlnAlaValVala:gaAlaAlaGlyGlyThrLeuLeuGlyValSerSer 214
 DB 241 GATAAGGAGTTCCAGGCGGTGGTCCGGGACGTGGTGTACCTCTGGGAGTCAGCTCC 300
 QY 215 GlnGlyGlyGlnThrTyrGluMetPheSerTrpPheLeuArgProLeuProGlyProHis 234
 DB 301 CAGGGGGGTGAGACCTACGAGATGTTCTCTGGTCTTCGGGCCCTCCAGGCCCCCAG 360
 QY 235 LysGlnLeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGlnGln 254
 DB 361 AGCAGCTCTCTCCACACGTCAGACCTTGCTGCTTCCAGTCCAGTCCGCGAGTGCAGCAG 420
 QY 255 HisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeu 274
 DB 421 CACCAGGGGAACCTGGATGCTTCGGGCCCGGACGTACCTTGTGAGGCTTCTCTGCTG 480
 QY 275 LysMetAlaGlnGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMet 294
 DB 481 AAGATGGCAGCAGGAGGAGCAAAACCCAGGCGACAGATTCCACACAGAACATGCTGATG 540
 QY 295 ThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyrThr 314
 DB 541 ACAGTCATTATTGCTTTGCTGGGACGATGACGGTCCAGCAGCAGGTCGGCTATACC 600

DB 123 CGGTGCTGGTGGAGACATTTCCA-GGGAGAGAGAGACGCCCATTCGATCCCTCCCTGCTG 181
 QY 174 LeuAlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyr 193
 DB 182 CTGGCCCGGACACTCCAAAGTAGTCTGCTCCCTCTCTTTGGCCTCGGCTTCTCTAT 241
 QY 194 GluAspLysGluPheGlnAlaValValArgAlaAlaGlyGlyThrLeuLeuGlyValSer 213
 DB 242 GAGGATAGGAGTTCCAGGCGGTGTCGGGCGAGCTGGTGTGTACCTCTCGGAGTCAGC 301
 QY 214 SerGlnGlyGlnThrTyrGluMetPheSerTrpPheLeuArgProLeuProGlyPro 233
 DB 302 TCCAGGGGGGTGACACCTACGAGATGTTCTCTGTTCTTCGGGCCCTGCCAGGCCCC 361
 QY 234 HisLysGlnLeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGln 253
 DB 362 CACAGCAGCTCTCCACACGTCAGCACCTTGGTGTGCTTCCAGTCCGCGAGTGCAG 421
 QY 254 GlnHisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAspAlaPheLeu 273
 DB 422 CAGCACCGGGAGACCTGGATGCTTTCGGGCCCGGACGTCGCTTGTGATGCCCTTCCTG 481
 QY 274 LeuLysMetAlaGlnGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeu 293
 DB 482 CTGAAGATGGCAGGAGGAGCAAAACCCAGGCGACAGATTCCACACAGAACATGCTG 541
 QY 294 MetThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyr 313
 DB 542 ATGACAGTCAATTATTGCTGTTGCTGGACGATGACGCTCAGCACACGCTCGGCTAT 601
 QY 314 ThrLeuLeuLeuMetLysTyrProHisValGlnLysTrpValArgGluGluLeuAsn 333
 DB 602 ACCCTCCTCTCTCGATGAATACCTCATGTGTCAAAGTGGGTACGTGAGGAGCTGAAT 561
 QY 334 ArgGluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThrArgLeuProTyrThr 353
 DB 662 CGGGAGCTGGGGCTGGCCAGGACCCAGCTAGGGGACCGTACCCGCTCCCTTACACC 721
 QY 354 AspAlaValLeuHisGluAlaGlnArgLeuLeuAlaLeuValProMetGlyIleProArg 373
 DB 722 GAGCGGT-CTGCATGAGGCGCAGCGTCTGCTGGCGTGTGTGCTTGGGGAATACCCCG 780
 QY 374 ThrLeuMetArgThrThrArgPheArgGlyTyrThrLeuProGlnGlyThrGluValPhe 393
 DB 781 AACCTCATGCGGACACCCGTTTCCGAGGGTACACCTG-CCCCAGGCCACGAGGCTTC 839
 QY 394 ProLeuLeuGly 397
 DB 840 CCCCTCTCGGT 851

RESULT 9
 BU543186
 LOCUS
 DEFINITION BU543186 877 bp mRNA linear EST 13-SEP-2002
 AGENCOURT_10338701 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6575194
 S', mRNA sequence.
 ACCESSION BU543186
 VERSION BU543186.1 GI:22853669
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DPF
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku, A., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-3216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
 URL: <http://fantom.gsc.riken.go.jp/>
 Location/Qualifiers

1. .1876

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:1810020016"

/db_xref="MGI:2389896"

/db_xref="taxon:10090"

/clone="1810020016"

/sex="male"

/tissue_type="pancreas"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="10 day old"

28. .1503

/note="unnamed protein product; cytochrome P450, 2b20 (MGD)|MGI:1202389, GB|X99715, evidence: BLASTN, 99%, match=1334)

putative"

/codon_start=1

/protein_id="BAC25748.1"

/db_xref="GI:26389536"

/translation="MEPSVLLALLVGLLLARGHPKSRGNFPFPGRPPLGLNL QMDRGLKSTQUREKQYDVTVHLGPRVPMVLCGTDIRALVQAEAFSGRTVA VEPFTKEGYVIFANGRWKTRFRFSLATMRDQMGKRSVEERIQEAAQCLVEELRKS QAPLDTFFLFCQTANICISVFGEFPEYTDROFLRLLELFYOTFSLISFSFQMPPE LFSQFLKYFPGAHRQISKNQELLDYIGHSVEKRAATLDPSVPDFIDILYLRKEK SNQTEPHONLMVSLSPFAGTSTSTTLRYGFLMLKYPHVAEKVKQKEDIVQIGS HRLTLDRTKWPTDVAIHIEIQFSDLIPIGVPHRTVOTMFRGYLLPKNTDVPYPI LSSALHDQYFQSPQSPFDPDANGALKKSEALFPSTGKRICLGESIAARNELFLFP TSILQNFVSVAHVAPKPIDITPKESGIGKIPPTYQICFLAR"

ORIGIN

Alignment Scores:	8.15e-125	Length:	1876
Score:	1299.50	Matches:	250
Percent Similarity:	68.79%	Conservative:	96
Best local Similarity:	49.70%	Mismatches:	144
Query Match:	49.69%	Indels:	13
DB:	11	Gaps:	3

US-10-669-693-2 (1-504) x AK028103 (1-1876)

9 LeuLeuAlaLeuAlaLeuLeuLeuLeuThrLeuAlaLeuSerGlyThrArgAla 28
 |||||
 : : : : :
 46 CTCCTCTTGTCTCTCTGTGGGCTTCTGTACTTACCGAGGGGACACCCAAAGTCC 105
 |||||
 : : : : :
 : : : : :

```

Qy 29 ArgGlyHisLeuProGlyProThrProLeuLeuGlyAsnLeuLeuGlnLeu 48
Db 106 CGTGCAACTCCACCAGGACCCCGTCCCTACCCCTCTTGGGAACTCTTGCAGATG 165
Qy 49 ArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerLysLysTyrGlyProValPhe 68
Db 166 GACAGAGAGCGCTCTCAAGTCTTTTATTTCAGCTTCGAGAAATATATGGAGATGTGTC 225
Qy 69 ThrIleTyrLeuGlyProTyrArgProValValLeuValGlyGlnGluAlaValArg 88
Db 226 ACAGTGCACCTGGGACCA---AGGCTGTGGTATGCTGTGGAAACAGACACCAATAAGG 282
Qy 89 GluAlaLeuGlyGlyGlnAlaGluPheSerGlyArgGlyThrValAlaMetLeuGln 108
Db 283 GAGGCTCTGTGGCCCAAGCGCTTCTCTGGCCGGGACAGATTGCTGTCTGTGAG 342
Qy 109 GlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTyrArgGlnLeuArg 128
Db 343 CCAACCTTCAAGAAATATGTTGATCTTTGGCAATGGGAACGTTGGAGACCTTCGT 402
Qy 129 LysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeu 148
Db 403 AGATTCTCTCGCCACCATGAGAGACTTGGGATGGGAAAGAGAGTGTGGAGGACGG 462
Qy 149 IleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPhe 168
Db 463 ATTGAGGAGAAAGCCCAATGTTTAGTGGAGAACTGCGGAAATCCCGAGGAGCCCTGTG 522
Qy 169 AspProSerLeuLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGly 188
Db 523 GACCCCACTTCCTCTTCAGTGCATCAGCCGCAATATATCTGCTCCATGTTGGTTGA 582
Qy 189 LeuArgPheSerTyrGluAspLysGluPheGlnAlaValValAlaGlyGlyThr 208
Db 583 GACCGTTTCAGTACACAGACCTCAGTCTTGGCCTGTGTGAGCTGTCTATCAGACC 642
Qy 209 LeuLeuGlyValSerSerGlnGlyGlyGlnThrTyrGluMetPheSerTyrPheLeuArg 228
Db 643 TTTTCACTATAAGCTCATTCTCCAGCCAGATGTTTGAGCTCTTCTGGCTCTCTCGAAG 702
Qy 229 ProLeuProGlyProHisLysGlnLeuLeuHisValSerThrLeuAlaAlaPheThr 248
Db 703 TACTTTCCTGGTCCCAACAGACAAATCTCAAAACCTGCAGAACTCTCGATACATT 762
Qy 249 ValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeu 268
Db 763 GGCCATAGTGTGGAGAGCACAGGCCACCTTGGACCCAGTGTCCA---CGAGACTTC 819
Qy 269 ValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAsnProGlyThrGluPheThr 288
Db 820 ATTGATATTTACCTTCTCGCATGGAGAGGAAAGTCCAACCCAGCACACGAGTTCAT 879
Qy 289 AsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSer 308
Db 880 CACCAAGACCTTCATGATGTGTGCTCTCTCTCTCTTCTTGTGGCCAGCAGACGAGC 939
Qy 309 ThrThrValGlyTyrThrLeuLeuLeuMetLysTyrProHisValGlnLysTyrVal 328
Db 940 ACCAGCTCCGCTATGGCTTCTCTCATGTCTCAAGTACCCCACTGTTCAGAGAAAGTC 999
Qy 329 ArgGluGlnLeuAsnArgGluLeuGlyValAlaGlyGlnAlaProSerLeuGlyAspArgThr 348
Db 1000 CAAAAGGAGATGTATGAGTATCGGTCTACCCGGTACCAACCCCTGTATGACCGGACC 1059
Qy 349 ArgLeuProTyrThrAspAlaValLeuHisGlnAlaGlnArgLeuLeuAlaLeuValPro 368
Db 1060 AAAATGCCATACATGATGAGTATCATCCAGAGATTCAGAGATTTTCAGATCTTATACCT 1119
Qy 369 MetGlyIleProArgThrLeuMetArgThrThrArgPheArgGlyTyrThrLeuProGln 388
Db 1120 ATTGGAGTGCCACACAGAGTGACCAAGATATCATATGTTCCGAGGGTACCTGCTCCCAAG 1179
Qy 389 GlyThrGluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePheLysHis 408

```

```

Db 1180 AACACTGAGGTGTACCCCATCTCAGTTCAGCTCTACATGATCCACAGTACTTTGAAACA 1239
Qy 409 ProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHisGlu 428
Db 1240 CCAGACAGTTTCAATCTCTGACCATCTCTGATGCAATGGGGCTACTGAAGAAAGTCAA 1299
Qy 429 AlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGlyLeuAlaLysAla 448
Db 1300 GCTTTTCTGCGCTTCTCAACAGGAAGGCATTTGCTTGTGTGAAGCATTTGCCCGCAAC 1359
Qy 449 GluLeuPheLeuPhePheThrThrIleLeuGlnAlaPheSerLeuGluSerProCysPro 468
Db 1360 GAGTGTGTTCTTTCTTCAGTCCATCTCTCAGAACATCTCTGTGGCAGCCATGTGCT 1419
Qy 469 ProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAlaPhe 488
Db 1420 CTTAAGGACATTTGACCTCACTCCCAAGAGAGTGTGTATTTGGAAAATACCTCCACGTAC 1479
Qy 489 GlnLeu-----GlnValArgProThrAspLeu 497
Db 1480 CAGATCTGCTTCTTGGCCGCTGATGGCTGAGGACAGACAGTGGCCCGCAGTACTGTTG 1539
Qy 498 HisSerThr 500
Db 1540 AGAATGACT 1548

RESULT 11
BO957100
LOCUS
DEFINITION
838 bp mRNA linear EST 21-AUG-2002
AGENCOURT 8777144 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384722
5', mRNA sequence.

ACCESSION
BO957100
VERSION
BO957100.1 GI:22372578
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 838)
NIH-MGC http://mgc.nci.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHC2580 row: d column: 03
High quality sequence stop: 671.
FEATURES
Location/Qualifiers
1..838
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6384722"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN

```

Alignment Scores:

Pred. No.: 2,67e-117 Length: 838
 Score: 1222.50 Matches: 248
 Percent Similarity: 95.02% Conservative: 0
 Best Local Similarity: 95.02% Mismatches: 10
 Query Match: 46.75% Indels: 4
 DB: 13 Gaps: 3

US-10-669-693-2 (1-504) x BQ957100 (1-838)

```

QY 115 GlyValPheSerAsnGlyGluArgTrpArgGlnLeuArgLysPheThrMetLeuAla 134
DB 2 GGGGTTTTCTTCCAC - GGGGAGCGGTGGAGCGACGCTGAGGAAGTTTACCATGCTTGCT 60

QY 135 LeuArgAspLeuGlyMetGlyLysArgGluGlyGluLeuIleGlnIleGluAlaArg 154
DB 61 CTGCGGGACCTGGGATGGGAAGCGAGGAGCGAGAGCTATCCAGCGGAGGCCGG 120

QY 155 CysLeuValGluThrPheGlnGlyThrGluGlyArgProPheAspProSerLeuLeuLeu 174
DB 121 TGTCTGTGGACATTCAGGGGACAGAGACGCCCATTCATCCCTCCCTGCTGCTG 180

QY 175 AlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGlyLeuArgPheSerTyrGlu 194
DB 181 GCCCAGGCGACCTCCACAGCTAGTCTGCTCCCTCCCTTTGGCTCCGCTTCTCTATGAG 240

QY 195 AspLysGluPheGlnAlaValAlaArgAlaAlaGlyGlyThrLeuLeuGlyValSerSer 214
DB 241 GATAGGAGGTCCAGGCCGTGTCGGGCGAGCTGGTGTACCTCTCGGAGTCAGCTCC 300

QY 215 GlnGlyGlyGlnThrTyrGluMetPheSerTrpPheLeuArgProLeuProGlyProHis 234
DB 301 CAGGGGGTTCAGACCTACGAGATGTTCTCTGTTCTTGGGCCCCCTCCAGGCCCCAC 360

QY 235 LysGlnLeuLeuHisHisValSerThrLeuAlaAlaPheThrValArgGlnValGlnGln 254
DB 361 AAGCAGCTCTCTCCACCGTCAGCACCTTGGCTGCTTTCACAGTCCGCGCAGGTGCAG 420

QY 255 HisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeuValAspAlaPheLeuLeu 274
DB 421 CACAGGGGAACCTGGATGCTTCGGGCCCCCGACGTGACCTTGTGATGCTCTCTGCTG 480

QY 275 LysMetAlaGlnGluGlnAsnProGlyThrGluPheThrAsnLysAsnMetLeuMet 294
DB 481 AAGATGGCACAGGAGGAACAAACCCAGGCACAGAAATTCACCAACAGAAATGCTGATG 540

QY 295 ThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSerThrThrValGlyTyrThr 314
DB 541 ACATCATTTATTGCTGTTTGTGGACGATGACGGTCAGCACCCGCTCGGCTATATACC 600

QY 315 LeuLeuLeuMetLysTyrProHisValGlnLysTrpValArgGluLeuLeuAsnArg 334
DB 601 CTCCTGCTCTGATGAATACCCCTCATGTCCTCAAAAGTGGGTACGTGAGGAGCTGAATCG 660

QY 335 GluLeuGlyAlaGlyGlnAlaProSerLeuGlyAspArgThrArgLeuProTyrThrAsp 354
DB 661 GAGCTGGGGTGGCCAGGCACCAAGCTAGGAGCAGCTACCCGCCCTCCCTTACCCGGA 720

QY 355 ---AlaValLeuHisGluAlaGlnArgLeu---LeuAlaLeuValProMetGly---Ile 371
DB 721 CGCGGGTCTTCCATGATGAGCGGCANCGGCTGGCTGGCGGCTGGTGGCCCATGGGAAATA 780

QY 372 Pro 372
DB 781 CCC 783

```

RESULT 12

BC022141
 LOCUS BC022141
 DEFINITION Mus musculus cDNA sequence BC034834, mRNA (cDNA clone IMAGE:5101144), containing frame-shift errors.
 ACCESSION BC022141
 VERSION BC022141.1 GI:18381021

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1944)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,S.I., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,F.L., Schett,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzyzinski,M.I., Skalska,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 1944)
 Strausberg,R.
 Direct Submission
 Submitted (24-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzyzinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeeadi, Jacqueline Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzyzinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeeadi, Jacqueline Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 57 Row: i Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

This clone has the following problem: frame shifted.

Location/Qualifiers

1. .1944

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:5101144"

/tissue type="Liver, normal. 5 month old male mouse."

/clone_lib="NCI CGAP_li9"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

FEATURES
source

ORIGIN

Alignment Scores:	1.13e-108	Length:	1944
Pred. No.:	1146.00	Matches:	230
Percent Similarity:	67.28%	Conservative:	101
Best Local Similarity:	46.75%	Mismatches:	155
Query Match:	43.82%	Indels:	8
DB:	11	Gaps:	3

US-10-669-693-2 (1-504) x BC022141 (1-1944)
Qy 7 TtpAlaLeuLeuLeuAlaLeuAlaLeuLeuLeuLeuLeuThrLeuAlaLeu----- 23
Db 21 TGGGTCTCCCTACTCTTGTCTCTC-CTGGTTTGGTGATGCCCTCAGCCTCTTTCAGTG 79
Qy 24 ---SerglyThrArgAlaArgGlyHisLeuProProGlyProThrProLeuProLeu 42
Db 80 TGGCAAAAAATCGCACATGGGGGAGAGATGCCCTCGGCCCACTCCCATTTCCCATATT 139
Qy 43 GlyAsnLeuLeuGlnLeuArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerLys 62
Db 140 GGGAAATATACTGCAATTGGACCTCAAAGACATCCCTGCATCGCTTTCGAAGCTGGCCAAA 199
Qy 63 LysTyrGlyProValPheThrIleTyrLeuGlyProTyrArgProValValLeuVal 82
Db 200 GAGTATGGCCCTGTTTACACTCTGTACTTTGGATCCTGG---CCTACCGTGGTCTCTACAT 256
Qy 83 GlyGlnGluAlaValArgGluAlaLeuGlyGlyGlnAlaGluGluPheSerGlyArgGly 102
Db 257 GCCTATGTTGGTGAAGAAAGCGCTACTCAATCAAGGTGATGAGTTCTTTGGGAGAGGA 316
Qy 103 ThrValAlaMetLeuGlnGlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGlu 122
Db 317 CCTCTGCCCATTTTCGAAGATAGCCAGAGGACATGGAATTGT--TTCAGTCAAGGAGAG 375
Qy 123 ArgTyrArgGlnLeuArgLysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLys 142
Db 376 AGTGTGAACCTCTCGCAGCGCTTCTCCCTCATGCGCTGAAGAACTTCGGAATGGGAAG 435
Qy 143 ArgGluGlyGluGluLeuLeuGlnAlaArgCysLeuValGluThrPheGlnGly 162
Db 436 AGAAGCCTAGAGGAGAGGTGCAGAGGAAGCCCGTGCTCGTGGAGAGTTACATAAA 495
Qy 163 ThrGluGlyArgProPheAspProSerLeuLeuLeuAlaGlnAlaThrSerAsnValVal 182
Db 496 ACAGAAGCTCAGCCCTTTGACCCCACTTCATCTTGGCTGTGCTCCCTGCAACGTGATC 555
Qy 183 CysSerLeuLeuPheGlyLeuArgPheSerTyrGluAspLysGluPheGlnAlaValVal 202
Db 556 TGCTCATCTCTTTCAACGAGCGATCCCATACATGACACACATTCCTCAACCTCATG 615
Qy 203 ArgAlaAlaGlyThrLeuLeuGlyValSerSerGlnGlyGlnThrTyrGluMet 222
Db 616 GACCTTTTAAATAAAAAATTTTACCACACTCTATATGGATCCAGATGTACACCTTA 675
Qy 223 PheSerTyrPheLeuArgProLeuProGlyProHisLysGlnLeuLeuHisValSer 242
Db 676 TGGCCGACGATCATGAAGTACATACCTGGAAGCATAGAGATTCTCAAAAAGGCTTGGT 735
Qy 243 ThrLeuAlaAlaPheThrValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSer 262
Db 736 GGTGTATAAAATCTTCACTCTAGAAAAAGTGAAGGAGCACCAGAGTCCCTGGACCCCGCC 795
Qy 263 GlyProAlaArgAspLeuValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAsn 282
Db 796 AACCCCT---CGGGATTATCATCGACTGTTTCTTCAGCAGATCGAGGAGAAACACAC 852
Qy 283 ProGlyThrGluPheThrAsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAla 302
Db 853 CTGAATCTGATTTTAACTCGAGAATTTGGCCATCTGTGGGTCTAACTTGTTTACGGCA 912
Qy 303 GlyThrMetThrValSerThrThrValGlyTyrThrThrLeuLeuLeuMetLysTyrPro 322


```

Db      1292 TACTTCATGACTTTCTCAGCAGGAAACGGGGTTTGTGCTGCTGAGAGGGCCTGGCCGGCATG 1351
Qy      449 GluLeuPheLeuPheThrThrIleLeuGlnIaPheSerLeuGluSerProCysPro 468
      |||||
Db      1352 GAGCTGTTCTTAATCCTGACCAACCATTTTACAGAACITTTAAACCTGAAATCTCTGTTGCAC 1411
Qy      469 ProAspThrLeuSerLeuLeuProThrValSerGlyLeuPheAsnIleProProAlaPhe 488
      |||||
Db      1412 CCAAAAGACATTGATATGATCCCATTTGTGAATGATTAATCACTCTGCCACCCCAITAC 1471
Qy      489 GlnLeu 490
      |||||
Db      1472 CAGCTC 1477

RESULT 15
AK008688
LOCUS   Mus musculus adult male stomach cDNA, RIKEN full-length enriched
DEFINITION library, clone:2210009K14 product:similar to CYTOCHROME P450 2C18
      (EC 1.14.14.1) (CYP1C18) (P450-6B/29C) [Homo sapiens], full insert
      sequence.
ACCESSION AK008688
VERSION   GI:12843035
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
  1 Carninci,P. and Hayashizaki,Y.
    TITLE High-efficiency full-length cDNA cloning
    JOURNAL Meth. Enzymol. 303, 19-44 (1999)
    MEDLINE 99279253
    PUBMED 10349636
REFERENCE
  2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
    Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
    TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
    prepare full-length cDNA libraries for rapid discovery of new genes
    JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
    MEDLINE 20499374
    PUBMED 11042159
REFERENCE
  3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
    Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
    Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
    Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
    Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
    Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
    Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
    TITLE RIKEN integrated sequence analysis (RISA) system--384-format
    sequencing pipeline with 384 multicapillary sequencer
    JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
    MEDLINE 20530913
    PUBMED 11076861
REFERENCE
  4 The RIKEN Genome Exploration Research Group Phase II Team and the
    FANTOM Consortium.
    TITLE Functional annotation of a full-length mouse cDNA collection
    JOURNAL Nature 409, 685-690 (2001)
    REFERENCE
  5 The FANTOM Consortium and the RIKEN Genome Exploration Research
    Group Phase I & II Team.
    TITLE Analysis of the mouse transcriptome based on functional annotation
    of 60,770 full-length cDNAs
    JOURNAL Nature 420, 563-573 (2002)
    REFERENCE
  6 (bases 1 to 1870)
    Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
    Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
    Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
    Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
    Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
    Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,

```

```

Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,K., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGAGATCAAGAGCTCTTTTITTTTNN 3'], cDNA was
prepared by using rehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGAGATTCCTCAGTTAATTAAATCAATCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOUR.

FEATURES             Location/Qualifiers
     source            1..1870
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2210009K14"
                     /db_xref="MGI:1905087"
                     /db_xref="taxon:10090"
                     /clone="2210009K14"
                     /sex="male"
                     /tissue_type="stomach"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS               1..1427
                     /note="unnamed protein product; putative
                     similar to CYTOCHROME P450 2C18 [EC 1.14.14.1] (CYP1C18)
                     (P450-6B/29C) [Homo sapiens] (SWISSPROT|P33260, evidence:
                     PASTY, 75%ID, 96.3%length, match=1416)"
                     /codon_start=3
                     /protein_id="BAB25834.1"
                     /db_xref="GI:12843036"
                     /translation="PKLWRONSQRNLPDGPPTPLIIGNILOLDKDISKSNFYSKV
                     YGPTTYLGRNPAVLHGYSYAEKAEFTDHEEAGRGVFFVDFKFKNCGVFSSGR
                     TWKERRPFLMTRNFGMRRSIEDRIQEAEARCLVDEURKTKGECPCDFTFLGASCNC
                     VICSIVQNFQDYKQDFELFDLIDNENVEISFLQICNNFNFAVIDLPGRRHLKH
                     KNFAEYFLSKVQKQESLIDINPRDFIDCFLLKMEQKHNKFTETCNLVFTAS
                     DLFAAGTETTTTLRYSLILLKYPKTEVAKVQEEIDCVIGHRSFCMDRHSMPYTD
                     VLHIOYIDLTLPSLPHAVTRDYKFERVLLPKGTIVIASLTSLVDYDKEPENPEKFD
                     PSHFDERGFKSKSDYFPFSTGKRICVGEGLARAELEFLTLTLQNFNLKSPVDLKD
                     LDITPVGANGFASVPPKFCICFIPI"
                     /note="putative"
                     1851..1856
                     /note="putative"
                     1870
                     /note="putative"

ORIGIN
Alignment Scores:
Pred. No.:      1,28e+100      Length:      1870
Score:          1069.00      Matches:      213
Percent Similarity: 64.94%      Conservative: 87
Best Local Similarity: 46.10%      Mismatches: 160
Query Match:      40.88%      Indels:      2
DB:              11      Gaps:      2
US-10-669-693-2 (1-504) x AK008688 (1-1870)

```

```
QY 29 ArgGlyHisLeuProProGlyProThrProLeuLeuGlyAsnLeuLeuGlnLeu 48
Db 30 AGAAGAAACCTCTCTCTGCCCCACACCTCTCTCCCATCATTTGAAATATTTCTAGCTA 89
QY 49 ArgProGlyAlaLeuTyrSerGlyLeuMetArgLeuSerLysLysTyrGlyProValPhe 68
Db 90 GATCTTAAGACATCAGCAATCTCTGAGGAATTTTCAAAAGTCTATNGCCCTGTGTTTC 149
QY 69 ThrIleTyrLeuGlyProTrpArgProValValLeuValGlnGlnAlaValArg 88
Db 150 ACTCTGTACTTGGGCAGG---AATCTCTGCTGTGTGTATCATGATGAAGCTGTGAAA 206
QY 89 GluAlaLeuGlyGlyGlnAlaGluPheSerGlyArgGlyThrValAlaMetLeuGlu 108
Db 207 GAAGCCTTTACTGATCATGCGGAGGAGTTTGTCTGGAAGAGAGTTTTCAGTGTGTGAT 266
QY 109 GlyThrPheAspGlyHisGlyValPhePheSerAsnGlyGluArgTrpArgGlnLeuArg 128
Db 267 AAGCTTTAAGAAACATCGCGGGTGTGTTTTCAGCAGTGGGAGGACCTTGGAAAGAGATGAGA 326
QY 129 LysPheThrMetLeuAlaLeuArgAspLeuGlyMetGlyLysArgGluGlyGlnLeu 148
Db 327 CGCTTCTCACTCACTCTCGCAAACTTTGGGATGGGTAGGAGGACATTGAGGACCGA 386
QY 149 IleGlnAlaGluAlaArgCysLeuValGluThrPheGlnGlyThrGluGlyArgProPhe 168
Db 387 ATTCAGAGAGGCTCTGCTCTGTGGATGAATTCAGGAAACCAAGGTGAACCTGT 446
QY 169 AspProSerLeuLeuAlaGlnAlaThrSerAsnValValCysSerLeuLeuPheGly 188
Db 447 GATCCCACTTCATCTCGCTGTGCTTCTTGTCAATGTCTATCTGCTCCATGTTTCCAG 506
QY 189 LeuArgPheSerTyrGluAspLysGluPheGlnAlaValValArgAlaGlyGlyThr 208
Db 507 AATCGTTTTGATTATAAAGATCAGGAATCTTACCTTCTTAGACATCTTTAAATGAAAT 566
QY 209 LeuGlyValSerSerGlnGlyGlyGlnThrTyrGluMetPheSerTrpPheLeuArg 228
Db 567 GTGGAGATCTGAGTTCCCTCGATTTCAGATTGCAATTAATTTTCTGCTGTCTATGAT 626
QY 229 ProLeuProGlyProHisLysGlnLeuLeuHisValSerThrLeuAlaAlaPheThr 248
Db 627 TATCTACCAAGGAGACACAGAAATACATAAAATTTTGCTTTGCAGAACATTACTTT 686
QY 249 ValArgGlnValGlnGlnHisGlnGlyAsnLeuAspAlaSerGlyProAlaArgAspLeu 268
Db 687 TTGTCAAAAGTAAACAGCATCAAGAGTCACCTGGATATTAAACAATCCT---CGGGACTTT 743
QY 269 ValAspAlaPheLeuLeuLysMetAlaGlnGluGlnAsnProGlyThrGluPheThr 288
Db 744 ATTGATTGTTTCTGATCAAAATGGAAACAGGAAACACACCCCAAGACGGAGTTTACT 803
QY 289 AsnLysAsnMetLeuMetThrValIleTyrLeuLeuPheAlaGlyThrMetThrValSer 308
Db 804 TGTGAGAACTTGGTCTTCACAGCATCTGATCTGTTGCTGGGACAGAGACAAAGC 863
QY 309 ThrThrValGlyTyrThrLeuLeuMetLysTyrProHisValGlnLysTrpVal 328
Db 864 ACAACCTGTAGGTACTCTCTGCTGCTCTGCTCAAGTATCCAGAACTCAGCCCAAGTC 923
QY 329 ArgGluGluLeuAsnArgGluLeuValAlaGlyGlnAlaProSerLeuGlyAspArgThr 348
Db 924 CAGGAAGAGATCGACTGTGTGATTGGCAGACAGAGGCCCTGTCATGCGAGCAGGCAT 983
QY 349 ArgLeuProTyrThrAspAlaValLeuHisGlnAlaGlnArgLeuLeuAlaValPro 368
Db 984 AGCATGCCCTACACTGATGCTGTGTGATGATCCAGATACATGATCTCTCTCC 1043
QY 369 MetGlyIleProArgThrLeuMetArgThrArgPheArgGlyTyrThrLeuProGln 388
Db 1044 ACCAGCCTTCCCATGCCGTTGACCCCTGAGCTGAAATTCAGAGAGTATCTCATTTCCCAAG 1103
```

```
QY 389 GlyThrCluValPheProLeuLeuGlySerIleLeuHisAspProAsnIlePheLysHis 408
Db 1104 GGCACCACCTGTCTATAGCATCACTACTCTGTCTATATGATAAAGAAATTTCCCAAC 1163
QY 409 ProGluGluPheAsnProAspArgPheLeuAspAlaAspGlyArgPheArgLysHisGlu 428
Db 1164 CCAGAGAAGTTTGACCCAGTCATCTTCTGATGAGCGTGGCAAGTTTAAGAAAAGTGAC 1223
QY 429 AlaPheLeuProPheSerLeuGlyLysArgValCysLeuGlyGluGlyLeuAlaLysAla 448
Db 1224 TACTTTTTTCTCTTCTCAACAGGAAAACGATTTGTGGAGAGGGTCTGGCTCGAGCA 1283
QY 449 GluLeuPheLeuPheThrThrThrLeuGlnAlaPheSerLeuGluSerProCysPro 468
Db 1284 GAATTATTTTATTCTGACCACCATTTTACAGAACTTTTAACCTTGAATCTCCAGTTGAT 1343
QY 469 ProAspThrLeuSerLeuLysProThrValSerGlyLeuPheAsnIleProProAlaPhe 488
Db 1344 CTAAAGACCTTGATACTACTCCAGTTGCCAATGGATTTCTTCTGTGCCCCCAAGTTC 1403
QY 489 GlnLeu 490
Db 1404 CAGATA 1409
```

Search completed: September 18, 2004, 04:33:32
Job time : 3799 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 17, 2004, 06:17:00 ; Search time 11362.9 Seconds
(without alignments)
11447.105 Million cell updates/sec
Title: US-10-669-693-3_COPY_1980_4980
Perfect score: 3001
Sequence: 1 cttttctctctccacc.....atcctggctaacacagttaa 3001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2997.8	99.9	129402	9	AC011510	AC011510 Homo sapi
c 2	2997.8	99.9	173251	2	AC058798	AC058798 Homo sapi
c 3	479.8	16.0	321519	2	AL714004	AL714004 Homo sapi
c 4	276.4	9.2	9532	9	HS999D10	Z94802 Human DNA s
5	272.2	9.1	126053	9	AL359916	AL359916 Human DNA
6	263.8	8.8	179607	2	AC027272	AC027272 Homo sapi
7	259.2	8.6	188833	9	HS268H5	AL008718 Human DNA
8	258	8.6	39634	9	AC139315	AC139315 Homo sapi
c 9	258	8.6	40948	2	AC140106	AC140106 Homo sapi
c 10	258	8.6	41867	2	AC144507	AC144507 Homo sapi
c 11	258	8.6	153609	9	HSJ9617	AJ009617 Homo sapi
c 12	258	8.6	158103	9	AC007952	AC007952 Homo sapi
c 13	258	8.6	158103	9	AC007952	AC007952 Homo sapi
c 14	258	8.6	176584	2	AC062010	AC062010 Homo sapi
15	257	8.6	79528	6	BD176843	BD176843 A method
16	257	8.6	79528	9	HS466N1	Z97630 Human DNA s
17	256.8	8.6	176221	2	AC011111	AC011111 Homo sapi
18	256.2	8.5	128611	9	AC083849	AC083849 Homo sapi
19	256	8.5	128624	9	AC008896	AC008896 Homo sapi
20	254.8	8.5	165735	9	AC135893	AC135893 Homo sapi
c 21	254.8	8.5	167110	2	AC022621	AC022621 Homo sapi
c 22	254.8	8.5	170269	2	AC144530	AC144530 Homo sapi
c 23	254.6	8.5	129036	9	AC005378	AC005378 Homo sapi
c 24	254.6	8.5	205642	9	AC016586	AC016586 Homo sapi
c 25	254.2	8.5	154179	9	AC087533	AC087533 Homo sapi
c 26	254.2	8.5	184511	2	AC011721	AC011721 Homo sapi
c 27	254.2	8.5	205638	9	AC009634	AC009634 Homo sapi
c 28	253	8.4	40351	9	AC005620	AC005620 Homo sapi
c 29	253	8.4	152945	9	AC098475	AC098475 Homo sapi
c 30	253	8.4	174539	9	AC103558	AC103558 Homo sapi
c 31	252.8	8.4	151999	2	AC005973	AC005973 Homo sapi
c 32	252.6	8.4	74155	9	AC004656	AC004656 Homo sapi
c 33	252	8.4	225370	9	AL445222	AL445222 Human DNA
c 34	251.8	8.4	63605	9	AL662787	AL662787 Human DNA
c 35	251.8	8.4	142916	9	AL928921	AL928921 Human DNA
36	251	8.4	171395	9	AC136605	AC136605 Homo sapi
c 37	251	8.4	179443	9	AC093290	AC093290 Homo sapi
38	251	8.4	182048	2	AC093293	AC093293 Homo sapi
39	251	8.4	186323	9	AC136601	AC136601 Homo sapi
c 40	251	8.4	262628	2	AC136609	AC136609 Homo sapi
c 41	250.8	8.4	2260	9	AY342392	AY342392 Homo sapi
42	250.8	8.4	162495	9	AC009996	AC009996 Homo sapi
c 43	250	8.3	69705	2	AC142832_3	Continuation (4 of
c 44	249.6	8.3	119043	9	AC139451	AC139451 Homo sapi
45	249	8.3	72539	9	AC119396	AC119396 Homo sapi

ALIGNMENTS

RESULT 1
AC011510
LOCUS AC011510 129402 bp DNA linear PRI 23-OCT-2000
DEFINITION Homo sapiens chromosome 19 clone CTD-2195B23, complete sequence.
AC011510
AC011510.7 GI:10947023
VERSION
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 129402)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished

REFERENCE
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT
On Oct 23, 2000 this sequence version replaced gi:9211205.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.7.

FEATURES
source
1. .129402
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2195B23"

ORIGIN
Query Match : 99.9%; Score 2997.8; DB 9; Length 129402;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CTTTCTCTCTCCCTACCCAGCTGAGTAAGTACGACCGGTGTTCCACCATCTACCT 60
DB 23073 CTTTCTCTCTCCCTACCCAGCTGAGTAAGTACGACCGGTGTTCCACCATCTACCT 23132
QY 61 GGGACCCCTGGCGCCTGTGTGCTCTGTTGGGAGGAGGTGTGCGGGAGGCCCTGGG 120
DB 23133 GGGACCGCTGGCGCCTGTGTGCTCTGTTGGGAGGAGGTGTGCGGGAGGCCCTGGG 23192
QY 121 AGGTGAGGCTGAGGAGTTGAGCGGCCGGGAAACCGGTAGGATGCTGGAAAGGACATTTGA 180
DB 23193 AGGTGAGGCTGAGGAGTTGAGCGGCCGGGAAACCGGTAGGATGCTGGAAAGGACATTTGA 23252
QY 181 TGGCCATGTAGTCAAGGCTGCTAGGCGCTCCGCTCACAGCTGCCACCACTTACTGG 240
DB 23253 TGGCCATGTAGTCAAGGCTGCTAGGCGCTCCGCTCACAGCTGCCACCACTTACTGG 23312
QY 241 TGTGTGACCTTTGCACATGGCTTAGTCCCTCTGTTGCTCTCATCTGTCAAATGGAGTGATA 300
DB 23313 TGTGTGACCTTTGCACATGGCTTAGTCCCTCTGTTGCTCTCATCTGTCAAATGGAGTGATA 23372
QY 301 ACAGTGCCCATGACCGGGTGCAGTGGCTAGTGGCTGAAATCCCAACACTTTGGGAGGCG 360
DB 23373 ACAGTGCCCATGACCGGGTGCAGTGGCTAGTGGCTGAAATCCCAACACTTTGGGAGGCG 23432
QY 361 GAGGTGGGTGGATCACTTGAGGTGAGAGTTGAGACCAAGCTGGCCCAACATGGTGAAC 420
DB 23433 GAGGTGGGTGGATCACTTGAGGTGAGAGTTGAGACCAAGCTGGCCCAACATGGTGAAC 23492
QY 421 CTGTCTCTACTAAAAATATAAAAAATTAGCTGGGCAATGGTGGCTACCTGTAATCCCA 480
DB 23493 CTGTCTCTACTAAAAATATAAAAAATTAGCTGGGCAATGGTGGCTACCTGTAATCCCA 23552
QY 481 GATACCTGGAGGTTGAGCGCGGAGTGCCTTGAACCCGGGAGGAGATGTTGCAGTGA 540
DB 23553 GATACCTGGAGGTTGAGCGCGGAGTGCCTTGAACCCGGGAGGAGATGTTGCAGTGA 23612
QY 541 ACCAGACTGTGCCACTGCAGTCCAGTCTGGGCAACAGAGTGAGCCCTCCATCTCAAACAA 600
DB 23613 ACCAGACTGTGCCACTGCAGTCTGGGCAACAGAGTGAGCCCTCCATCTCAAACAA 23672
QY 601 ACAACAAAAAGCAGTGCCCATCATGTAGATTTAGTGGAGTATGAGTGAGGACTGAGCCTTG 660
DB 23673 ACAACAAAAAGCAGTGCCCATCATGTAGATTTAGTGGAGTATGAGTGAGGACTGAGCCTTG 23732

QY 661 TCGAAGTGAGCACTCAGTAAATCACCAGGTTGTAGTATCAGTGATACCATCAATGATCC 720
DB 23733 TCGAAGTGAGCACTCAGTAAATCACCAGGTTGTAGTATCAGTGATACCATCAATGATCC 23792
QY 721 AGGTAAAGCCCTGAGGTTTCAGAAAGATCCCGAGCGCTTTCAAGTGTCTGGGATTTGGT 780
DB 23793 AGGTAAAGCCCTGAGGTTTCAGAAAGATCCCGAGCGCTTTCAAGTGTCTGGGATTTGGT 23852
QY 781 GGGCAAGCCCTCGAATAATAGAAAACAGTTCTCTGTATTACAAACAGAAAACAGGAGGCCCA 840
DB 23853 GGGCAAGCCCTCGAATAATAGAAAACAGTTCTCTGTATTACAAACAGAAAACAGGAGGCCCA 23912
QY 841 TCGTGGTGTCTGCCAGGAACCTCAGTAGTAACCTAAGACACACCGGTGCTGTCTCCCCAGC 900
DB 23913 TCGTGGTGTCTGCCAGGAACCTCAGTAGTAACCTAAGACACACCGGTGCTGTCTCCCCAGC 23972
QY 901 GCACCTAGGCCAGTGGGAAACAGACTCACACACAGTCCCAGCCCCAGAGTGGTCAGGGC 960
DB 23973 GCACCTAGGCCAGTGGGAAACAGACTCACACACAGTCCCAGCCCCAGAGTGGTCAGGGC 24032
QY 961 CAAGATGGGAAACAGCGGGAGAAAGGTCAGGGTGGGATGGGAGGGGTGAGGGCAAGAG 1020
DB 24033 CAAGATGGGAAACAGCGGGAGAAAGGTCAGGGTGGGATGGGAGGGGTGAGGGCAAGAG 24092
QY 1021 GGGTCAGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGAC 1080
DB 24093 GGGTCAGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGAC 24152
QY 1081 CCGGCATGTTTGGTGGGAGATTCAGGAAGTCTTCTCGAAGAGAGGCTGTGCGAGCTG 1140
DB 24153 CCGGCATGTTTGGTGGGAGGCTTCAGGAAGTCTTCTCGAAGAGGCTGTGCGAGCTG 24212
QY 1141 AGACTCATAAGATGATGGGAGGTTTCCAGGCAGAAAGACCAAGCACTCAAAAGCA 1200
DB 24213 AGACTCATAAGATGATGGGAGGTTTCCAGGCAGAAAGACCAAGCACTCAAAAGCA 24272
QY 1201 TGACCTTGAGCAAGCAATTCATCACTCACTGATGAATTTTCAGCTGGGCAAGCTGGC 1260
DB 24273 TGACCTTGAGCAAGCAATTCATCACTCACTGATGAATTTTCAGCTGGGCAAGCTGGC 24332
QY 1261 TCATGCTCTTAATCCAGCACTTTGGAAGGCTGAATGGGAGGATGACTTTAGGCTTAGGC 1320
DB 24333 TCATGCTCTTAATCCAGCACTTTGGAAGGCTGAATGGGAGGATGACTTTAGGCTTAGGC 24392
QY 1321 ATTGTGACAGCCTGGGCAACATGGTGAGACCTGCTCCACAAAACAAAACAAAC 1380
DB 24393 ATTGTGACAGCCTGGGCAACATGGTGAGACCTGCTCCACAAAACAAAACAAAC 24452
QY 1381 AAAAAATCAATTATACCTGTTACCATGGGTACCAAGGTACATAGAAATGACTCAGGCAGATA 1440
DB 24453 AAAAAATCAATTATACCTGTTACCATGGGTACCAAGGTACATAGAAATGACTCAGGCAGATA 24512
QY 1441 TGGTGTCTCTCTCTACTGTGGGAGGCGGGCTTATCTGCAGTAAACAATAAGAGGGAG 1500
DB 24513 TGGTGTCTCTCTCTACTGTGGGAGGCGGGCTTATCTGCAGTAAACAATAAGAGGGAG 24572
QY 1501 GGAATATAATCTCTAAATAGAGGTACAGATTTGAGACAAACACAGGGCACAGGCATAT 1560
DB 24573 GGAATATAATCTCTAAATAGAGGTACAGATTTGAGACAAACACAGGGCACAGGCATAT 24632
QY 1561 GTACGAGGTTAAAGAGGGAATCAGGGAAGCTTCTCAGAGAGGTGACATTTAAAGCCGG 1620
DB 24633 GTACGAGGTTAAAGAGGGAATCAGGGAAGCTTCTCAGAGAGGTGACATTTAAAGCCGG 24692
QY 1621 ACATGAAGGATGAACAGGTTAGTTTCAACCAAGATGGATGGAAAGGGGTGAGAGTATG 1680
DB 24693 ACATGAAGGATGAACAGGTTAGTTTCAACCAAGATGGATGGAAAGGGGTGAGAGTATG 24752
QY 1681 AGGCAAGGGAACTGAGGATCATAGGCTTAGAGGGGATCTCTGACGCGCTTTGAGGAAG 1740
DB 24753 AGGCAAGGGAACTGAGGATCATAGGCTTAGAGGGGATCTCTGACGCGCTTTGAGGAAG 24812

QY 1741 TGAGAGAACACAGCGAGTCGTAGTGGTTAAAGTAACAAAGCTGAGAGCAGGGAAT 1800
Db 24813 TGAGAGAGACAGCGCAGTCGTAGTGGTTAAAGTAACAAAGCTGAGAGCAGGGAAT 24872
QY 1801 CCCTGGTCATCAGAGGCTGTGAGTCAGCTCAGAGTGTTCGGGCTTTTGTTCCTGGG 1860
Db 24873 CCCTGGTCATCAGAGGCTGTGAGTCAGCTCAGAGTGTTCGGGCTTTTGTTCCTGGG 24932
QY 1861 AGCAGTCGATTTAAAGCAGGGAACAGCTGTATTACAGAGTGGGAAGATCCTGTGGTCT 1920
Db 24933 AGCAGTCGATTTAAAGCAGGGAACAGCTGTATTACAGAGTGGGAAGATCCTGTGGTCT 24992
QY 1921 GCCTGAAGGGGATGAACCTGAGGCTAGAGCCAGGCTGATAGGAGGATCCAGGGTGA 1980
Db 24993 GCCTGAAGGGGATGAACCTGAGGCTAGAGCCAGGCTGATAGGAGGATCCAGGGTGA 25052
QY 1981 TGGGGAGCTGGGAGGTCGCGGTGATGGACAGGGCTGGGGCCAGGGGATGGGAGGAA 2040
Db 25053 TGGGGAGCTGGGAGGTCGCGGTGATGGACAGGGCTGGGGCCAGGGGATGGGAGGAA 25112
QY 2041 GGAGTAATTGGAGAGGCTGGGCTCTGGCGGAGGAATGATGGTGGGCTGAACAGGG 2100
Db 25113 GGAGTAATTGGAGAGGCTGGGCTCTGGCGGAGGAATGATGGTGGGCTGAACAGGG 25172
QY 2101 AGAGGAGATGCTTAGGCCACTTTGGAAACAGTAGGGCAAGGACAGGACACCCCAAG 2160
Db 25173 AGAGGAGATGCTTAGGCCACTTTGGAAACAGTAGGGCAAGGACAGGACACCCCAAG 25232
QY 2161 GGGAGGTGCCAAGAGACCCAGCAGCGCTGGCATTTGGACAGGGAGGTCCTGTGGAGCA 2220
Db 25233 GGGAGGTGCCAAGAGACCCAGCAGCGCTGGCATTTGGACAGGGAGGTCCTGTGGAGCA 25292
QY 2221 GGTCTCTTGGTAAGGAGGAAATGGTGAGTTCATCTCCCTCTCTGTTCACCC 2280
Db 25293 GGTCTCTTGGTAAGGAGGAAATGGTGAGTTCATCTCCCTCTCTGTTCACCC 25352
QY 2281 TCTAAACTACATGGGGCAGCAGACCCAGTGGGACTCCCAATATGATGGGATGGGTGGATG 2340
Db 25353 TCTAAACTACATGGGGCAGCAGACCCAGTGGGACTCCCAATATGATGGGATGGGTGGATG 25412
QY 2341 GAAGGAAGGAAGGAGGAACAACTCTTCATTCCTCGTTTATTATACAGAACAGGCCAGG 2400
Db 25413 GAAGGAAGGAAGGAGGAACAACTCTTCATTCCTCGTTTATTATACAGAACAGGCCAGG 25472
QY 2401 TGCGGTCTCAGCTTGCCTTCTAGACCTTTGGAGCTGAGGTGGGTGGATTAACCTCA 2460
Db 25473 TGCGGTCTCAGCTTGCCTTCTAGACCTTTGGAGCTGAGGTGGGTGGATTAACCTCA 25532
QY 2461 GGTGAGGATTCAGAGCAGGCTAGACAACTGAGAAACCCCATCTCTACTGAAGATAT 2520
Db 25533 GGTGAGGATTCAGAGCAGGCTAGACAACTGAGAAACCCCATCTCTACTGAAGATAT 25592
QY 2521 AAAATTAGCTGGCGTAGTGCATATGCTGTATCCAGCTAGTCGGGAAGCTGAGGCA 2580
Db 25593 AAAATTAGCTGGCGTAGTGCATATGCTGTATCCAGCTAGTCGGGAAGCTGAGGCA 25652
QY 2581 GGAGAACTGCTTGAACCCGAGCAGCAGAGGTTCGGTGAGCTGAGATGCTGCCATTGCAC 2640
Db 25653 GGAGAACTGCTTGAACCCGAGCAGCAGAGGTTCGGTGAGCTGAGATGCTGCCATTGCAC 25712
QY 2641 TCCAGCTGGGTGACAAAGCAAGACCTGCTCTCAATATAATAATAATTACAAAACAGAA 2700
Db 25713 TCCAGCTGGGTGACAAAGCAAGACCTGCTCTCAATATAATAATAATTACAAAACAGAA 25772
QY 2701 GGAGCTGGGTGATCCCAAGCTACCTTACTTTTTCAGGAGAAATGACTCCCTTACCCAAAGGGC 2760
Db 25773 GGAGCTGGGTGATCCCAAGCTACCTTACTTTTTCAGGAGAAATGACTCCCTTACCCAAAGGGC 25832
QY 2761 AAAGGATGGGAACCAAGTTTATATGATATGATATGATATGATGAGACCTACTGAGTCTCATCCC 2820
Db 25833 AAAGGATGGGAACCAAGTTTATATGATATGATATGATATGAGACCTACTGAGTCTCATCCC 25892
QY 2821 TGGGCTAGGCTGGAATGGACTCAGATGGAGCTGAAGAGTCCCTCTCAGGGAACTCTCACT 2880

Db 25893 TGGGCTAGGCTGGAAATGGAATCAGATGGAGCTGAAGAGTCCCTCAGGGAACTCACT 25952
QY 2881 AGAAGAAGAGGAGGAATCGGCCGGCGGGTGGCTCACGCCCTGTAAATCCCAACACTTTGGG 2940
Db 25953 AGAAGAAGAGGAGGAATCGGCCGGCGGGTGGCTCACGCCCTGTAAATCCCAACACTTTGGG 26012
QY 2941 AGGCTGAGGTGGGTGGATCAAGAGTCAGGAGATCGAGACCATCTCTGCTTAACAGTGA 3000
Db 26013 AGGCTGAGGTGGGTGGATCAAGAGTCAGGAGATCGAGACCATCTCTGCTTAACAGTGA 26072
QY 3001 A 3001
Db 26073 A 26073
RESULT 2
AC058798/c
LOCUS AC058798 173251 bp DNA linear HTG 19-APR-2000
DEFINITION Homo sapiens clone RP11-166H18, WORKING DRAFT SEQUENCE, 35
unordered pieces.
AC058798
VERSION AC058798.1 GI:7596893
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 173251)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens, clone RP11-166H18
Unpublished
2 (bases 1 to 173251)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lechoczky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mleaga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tefaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (19-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L7830

Center clone name: 166_H18

Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156916 bases at least Q40

Consensus quality: 164898 bases at least Q30
Consensus quality: 167842 bases at least Q20
Insert size: 178000; agarose-fp
Quality coverage: 3.9 in Q20 bases; agarose-fp
Quality coverage: 4.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 35 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1125: contig of 1125 bp in length
* 1126: gap of 100 bp
* 1226: contig of 1262 bp in length
* 2488: gap of 100 bp
* 2588: contig of 1698 bp in length
* 4286: gap of 100 bp
* 5474: contig of 1088 bp in length
* 5574: gap of 100 bp
* 6859: contig of 1286 bp in length
* 6959: gap of 100 bp
* 8244: contig of 1285 bp in length
* 8344: gap of 100 bp
* 8345: contig of 1171 bp in length
* 9515: gap of 100 bp
* 9516: gap of 100 bp
* 11084: contig of 1469 bp in length
* 11184: gap of 100 bp
* 11885: contig of 1653 bp in length
* 12838: gap of 100 bp
* 12938: contig of 1884 bp in length
* 14822: gap of 100 bp
* 14822: contig of 2931 bp in length
* 17852: gap of 100 bp
* 17853: contig of 1962 bp in length
* 19915: gap of 100 bp
* 20015: contig of 2350 bp in length
* 22365: gap of 100 bp
* 22465: contig of 1706 bp in length
* 24171: gap of 100 bp
* 24271: contig of 4027 bp in length
* 28298: gap of 100 bp
* 28398: contig of 2851 bp in length
* 31349: gap of 100 bp
* 31349: contig of 2322 bp in length
* 33671: gap of 100 bp
* 33771: contig of 2850 bp in length
* 36821: gap of 100 bp
* 36821: contig of 2839 bp in length
* 39560: gap of 100 bp
* 39560: contig of 3714 bp in length
* 43374: gap of 100 bp
* 43474: contig of 2767 bp in length
* 46241: gap of 100 bp
* 46341: contig of 3813 bp in length
* 50154: gap of 100 bp
* 50254: contig of 3028 bp in length
* 53281: gap of 100 bp
* 53282: contig of 2999 bp in length
* 56381: gap of 100 bp
* 56381: contig of 4658 bp in length
* 61239: gap of 100 bp
* 61239: contig of 5643 bp in length
* 66882: gap of 100 bp
* 66882: contig of 8409 bp in length
* 75391: gap of 100 bp
* 75391: contig of 7442 bp in length
* 82932: gap of 100 bp
* 83033: contig of 10869 bp in length
* 93901: gap of 100 bp
* 93902

* 94002 106935: contig of 12934 bp in length
* 107035: gap of 100 bp
* 11729: contig of 10694 bp in length
* 11730: gap of 100 bp
* 11730: contig of 8284 bp in length
* 126113: gap of 100 bp
* 126114: contig of 12642 bp in length
* 138855: gap of 100 bp
* 138856: contig of 18325 bp in length
* 157281: gap of 100 bp
* 157281: contig of 15871 bp in length.
* 157381
FEATURES
Location/Qualifiers
1..173251
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-166H18"
/clone_lib="RPCI-11 Human Male BAC"
1..1125
misc_feature
1226..2487
/note="assembly_fragment"
2588..4285
/note="assembly_fragment"
4386..5473
/note="assembly_fragment"
5574..6859
/note="assembly_fragment"
6960..8244
/note="assembly_fragment"
8345..9515
/note="assembly_fragment"
9616..11084
/note="assembly_fragment"
11185..12837
/note="assembly_fragment"
12938..14821
/note="assembly_fragment"
14922..17852
/note="assembly_fragment"
17953..19914
/note="assembly_fragment"
20015..22364
/note="assembly_fragment"
22465..24170
/note="assembly_fragment"
24271..28297
/note="assembly_fragment"
28398..31248
/note="assembly_fragment"
31349..33670
/note="assembly_fragment"
33771..36620
/note="assembly_fragment"
36721..39559
/note="assembly_fragment"
39660..43373
/note="assembly_fragment"
43474..46240
/note="assembly_fragment"
clone_end:SP6
vector_side:right
46341..50153
/note="assembly_fragment"
50254..53281
/note="assembly_fragment"
clone_end:R7
vector_side:left
53382..56380
/note="assembly_fragment"
56481..61138
/note="assembly_fragment"

Query Match									
Best Local Similarity 99.9%; Score 2997.8; DB 2; Length 173251;									
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	1	CTTTCTTCTCCTACCCACCTGAGTGAAGATACGGACCGGTGTTTCCACCATCTACCT	60						
DB	123051	CTTTCTTCTCCTACCCACCTGAGTGAAGATACGGACCGGTGTTTCCACCATCTACCT	122992						
QY	61	GGGACCCCTGGCGGCTGTGTGCTCTGTTGGGAGGAGGCTGTGCGGAGGCGCTGGG	120						
DB	122991	GGGACCCCTGGCGGCTGTGTGCTCTGTTGGGAGGAGGCTGTGCGGAGGCGCTGGG	122932						
QY	121	AGGTGAGGCTGAGGAGTTTACGGGCGGGGAAACCGTAGCGATGCTTGGAGGACCTTTTGA	180						
DB	122931	AGGTGAGGCTGAGGAGTTTACGGGCGGGGAAACCGTAGCGATGCTTGGAGGACCTTTTGA	122872						
QY	181	TGGCCATGTTAAGTCAAGGGCTGTAGGCGCTCCCTCACAGCCTGCCACCACTTTACTGG	240						
DB	122871	TGGCCATGTTAAGTCAAGGGCTGTAGGCGCTCCCTCACAGCCTGCCACCACTTTACTGG	122812						
QY	241	TGTGTGACCTTGCACATGGCTTAGTCCCTCTGTTGGCTCATCTGTGCAATGGAGTGATA	300						
DB	122811	TGTGTGACCTTGCACATGGCTTAGTCCCTCTGTTGGCTCATCTGTGCAATGGAGTGATA	122752						
QY	301	ACAGTGGCCATCAGCCGGGTGCAAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCG	360						
DB	122751	ACAGTGGCCATCAGCCGGGTGCAAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCG	122692						
QY	361	GAGTGGGTGATCACTTGAGGTGAGGATCGAGATTCGAGACAGCCTGGCCAACTGTGAAAC	420						
DB	122691	GAGTGGGTGATCACTTGAGGTGAGGATTCGAGACAGCCTGGCCAACTGTGAAAC	122632						
QY	421	CCTGTCTCTACTAAATAATATAAATAATAGCTGGGATGCTGGTGGCTGATATCCCA	480						
DB	122631	CCTGTCTCTACTAAATAATATAAATAATAGCTGGGATGCTGGTGGCTGATATCCCA	122572						
QY	481	GATATCTTGGGAGGTGAGGACAGGAGATCGCTTGAACCCCGGAGGACAGATGTCAGTGA	540						
DB	122571	GATATCTTGGGAGGTGAGGACAGGAGATCGCTTGAACCCCGGAGGACAGATGTCAGTGA	122512						
QY	541	ACCAAGACTGTGCCACTGCACTCCAGTCTGGGCAACAGATGAGCCTCCATCTCAAAACA	600						
DB	122511	ACCAAGACTGTGCCACTGCACTCCAGTCTGGGCAACAGATGAGCCTCCATCTCAAAACA	122452						
QY	601	ACAAACAAAAGACGTGCCATCATGTAGGATTCAGTGTGATGAGTGGGATGAGTGGT	660						
DB	122451	ACAAACAAAAGACGTGCCATCATGTAGGATTCAGTGTGATGAGTGGGATGAGTGGT	122392						
QY	661	TGCAAGTGTGACACTCACTAATCAACCGAGTTGTAGTATCAGTGTGATAACCATGATGCC	720						
DB	122391	TGCAAGTGTGACACTCACTAATCAACCGAGTTGTAGTATCAGTGTGATAACCATGATGCC	122332						
QY	721	AGGTAAAGCCCTGAGGCTTCAAGAGATGCGGAGCGCTTCAAGGTGCTGGGATTTGGT	780						
DB	122331	AGGTAAAGCCCTGAGGCTTCAAGAGATGCGGAGCGCTTCAAGGTGCTGGGATTTGGT	122272						
QY	781	GGGCAAGCCCTCGAATAATAAGAACAGTCTCTCTATTACACAGAGAGGAGGCGCCA	840						
DB	122271	GGGCAAGCCCTCGAATAATAAGAACAGTCTCTCTATTACACAGAGAGGAGGCGCCA	122212						
QY	841	TGCTGGGTGTGTCAGGAACTCAGTAGTAATAAGACAGCACCGTCTCTCTCCCGCAGC	900						
DB	122211	TGCTGGGTGTGTCAGGAACTCAGTAGTAATAAGACAGCACCGTCTCTCTCCCGCAGC	122152						
QY	901	GCACCTTAGGCGAGTGGGAAACAGATCAACACAGTCCCGACCCAGAGTGTTCAGGGC	960						
DB	122151	GCACCTTAGGCGAGTGGGAAACAGATCAACACAGTCCCGACCCAGAGTGTTCAGGGC	122092						
QY	961	CAAGATGGGGAAGACACGGGAGAGAGGTTCAGGCTGGGATGGGAGGCTCAGGGCAAGAG	1020						
DB	122091	CAAGATGGGGAAGACACGGGAGAGAGGTTCAGGCTGGGATGGGAGGCTCAGGGCAAGAG	122032						
QY	1021	GGGTGAGGCGCAGGCTCAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGAC	1080						

Db 1211 CAGGGAGAAATTCCTTGAATCTGGAGGTGGAGGTTGCACTGAGCTGAGATTATGCCACTG 1270
QY 1195 AAAGCATGACCTTGAGAGAGCAATTCATCCATTCACTGATGAATTTTCAGACTGGGCAC 1254
Db 1271 CACTCCAGCCCTCAACGACAGAGCAAGACTCCATCTCCAAAAATAAAAAAGAACACAGGTGC 1330
QY 1255 GCTGCTCATGCTGTAATCCAGACACTTTGGAGGCTGAAATGGGGAGGATGACTTGAGC 1314
Db 1331 AGTGGCTCAGCCCTGTAATCCAGGACCTTTGGAGGCTGAGGCCGCGAGATCACCTGAGG 1390
QY 1315 CTAGGCATTTGTGACAAAGCTGGGCAACATGGTGAGACCCCTGCTCCACAAAAACAAACAA 1374
Db 1391 TCAGGAGTTCAAGACCAAGCCCTGGCCACATGGTGAACCCCATCTCTACTAAATACAA 1450
QY 1375 ACAAACAAAAAATCATATATACCTGTAGTACCATGGGTACAGGTACATAGAAATGACTCAGG 1434
Db 1451 AAAATTAGCTGGCGTGATGGGGTGCTGCTGTAATCCAGGCTACTCGGGGGGCTGAGAGC 1510
QY 1435 CAGATATGGTGT---CCTCTCTACTGTGGGAGAGCGGGCTTATPACTGCACTAGACAA 1491
Db 1511 GGAGAATTGCTTGAGCCCTGGGAGACGGTGGTTCAGTAAGTGAATGGATTCGACCATGCTG 1570
QY 1492 TAGAGGAGGGAATTAATCCTAAATGAGAGGTACAGATTGAGAGCAAAACACAGGSCA 1551
Db 1571 TCCAGCCCTGGGACGCTGAGCAAGACTCCATCTCAAAAAAAGAAAAAAGAAAGGCT 1630
QY 1552 CAGGCATATGTACAGAGGTAAAGAGGGAATCAGGAGGCTTCTCAGAAAGGTGACATT 1611
Db 1631 GGGTGAGTGGCTCATGCTTAATAATCCAGCAGCTTTGGGAGGCGGAGACGGGTGGATCAT 1690
QY 1612 TAAGCCGGGACATGAAGGATGAACGAGTTA-----GTTCAACCAAGGATGGATGMAAGG 1666
Db 1691 GAGTCAAGGAGAGGAGACCATCTGGCTACATGTTGTAACCTGCTCTACTATAAAAT 1750
QY 1667 GGTGAGAGTATGAGGAGGAGGAGAACTGCAAGATCATPAGCCCTPAGACAGGGGCTCTGA 1726
Db 1751 ACAAAAAAATTAGCCGGCGTGGTGGCGGCGCTATAGTCCAGCTACTCGGGAGGCTGA 1810
QY 1727 CGCCCTTGAGGAAGTGAAGAGACCAAGCAGCTGCTAGTGGTAAAGTAACAAAGCTGA 1786
Db 1811 GGC-----AGGAGATGGCTGAACCCGGGAGGAGAGCTTGCACTGAGGCAAGATCGCG 1865
QY 1787 GAAGCCAGGGAATPCCCTGCTCATGCAAGGCGCTGTGAGTCAGCTCAGCTGTTGGGCTT 1846
Db 1866 CCATTGCACTCCAGCCTGGCCACAGAGCGAACTCCATCTCAAAAAAAGAAAAA 1925
QY 1847 TTGTTTTCTGGAGCAGTCGATTTTAAGCAGGGAACAGCTGTATTTCAGAGTTGGGAAG 1906
Db 1926 AAGAACAAGTTGGAGGACTCTCATTTGTCAATTTAAATACTTACCACAAAGCTACATA 1985
QY 1907 ATCCTGTGGTTGCTGTGAAGGGATGAACCTGGAGGCTAGAGGCCAGGGGTGATAGG 1966
Db 1986 AT-----CAAAACAGCATGGCAGGCATGGTGGCTCACCTGTAATCCCA 2031
QY 1967 AGGATCAGGCTGATGGGAGGCTGGGAGGCTCGCGGTGATGACACAGGCTGGGGCCAG 2026
Db 2032 GCACCTTTGGAGGCTGAGGAGGCTGATCCTGAGGT-----CAGGAGTTCAAGACC 2084
QY 2027 GGGATGGGAGGAGAGTAATTGGGAGAGGCTGGGGCTCTGGCCGAGGAATGGATGGT 2086
Db 2085 AGCCTGGCCAAACATGTTGAACTCCATCTCTACTAAAAATACAAAAATTAGCTGGGTG 2144
QY 2087 GGCTCAAAACAGGAGGAGGAGATGCTTAGGCCACTTTGG--AACACAGTAGGGCAAGGA 2145
Db 2145 GTGGCAAAAGCCTGTATATCCAGCTACTCGGGAGGCTGAGGAGGAAATGCTTGAGCC 2204
QY 2146 CAGGAGACCCAAAGGGGAAGTCCCAAGAGACACAGCAGGCTGCTGATTTGACAGGAA 2205
Db 2205 CGGAGGAGAGGTTCCAGTGCATGAGTACGACACTGCACTCCAGGCTGACTGACAG 2264
QY 2206 GGTCTCTGGAGCAGGCTGCTTTGGTAATAGGGAGGAAATGGTGCAGTTCCTCCCTCCTC 2265

Db 2265 AGTGAGACTGTCTCAAAAACAAAAACAAAAACAAAAACAGTGTGGCACTGGCTTTAAAGACAG 2324
QY 2266 CTTCTCTGTTCACCTCTAAACTACATGGGGCACAGGACCCAGTGGGACTTCATTAATGA 2325
Db 2325 ACACAGAGATCAAT---GGGGTTAAATAGACAGGCCAGAAATAAACACTTCAATATATGG 2381
QY 2326 TGGGATGGTGTGATGGAGGAGGAGGAGGAGAAACAACTTTCAATTCATCTCTGGTTATTT 2385
Db 2382 TCAATGGTTTTTCAACAAGGGTGCAGAAATCATTTCAATGGTGAAGAATGATCTTTTCA 2441
QY 2386 ACAGAACAGGCGAGCTGGGT---GCTCAGCTTGCATTTCTAGCACTTTGGGAGGCTGAGG 2444
Db 2442 ACAATGGGGCAGAGTGGTGGCTCAGCCCTGTAATCTTAGCACTCTGGGAGGCGGAGG 2501
QY 2445 TGGGTGATTAACCTCAGGTCAGGAGTTCAAGACCAAGCCCTAGACACAGCTAGAGAAACCCCA 2504
Db 2502 TGGGCACATCATCTGAGTCAAGGTTTAAGACCAAGCTGCCAACAGTGGTGAATCTCTG 2561
QY 2505 TCTCTACTGAAGATATAAA---ATTAGCTGGGGCTGATGGCAATGCTCTTAATCCAGCTA 2563
Db 2562 TCTTGACTGAAAAATACAAACATTAGCTGGGCATGGTGGTGCATGCTGTAATTCAGCTA 2621
QY 2564 GTCCGGAGCTGAGCGAGGAGATCGCTTGAAACCGGAGGAGAGGTTGCGGTGAGCTG 2623
Db 2622 CTGAGAGGCTGAGCGAGGAGATCATCTGAGCCCGGAGGAGGTTGCACTGAGCGG 2681
QY 2624 AGATCGTGCCATTGCTCCAGCCCTGGGTGACAAAGCAAGACCTCGTCTCAATTAATA 2683
Db 2682 AGATCTGCCACTGCACTCAAGCTTGACACACAGCAGCACTCCATCTCAAAAAACAA 2741
QY 2684 ATANTTACAA 2693
Db 2742 CAAGCAAAA 2751

RESULT 5
AL359916
LOCUS

DEFINITION
AL359916 126053 bp DNA linear PRI 12-FEB-2001
Human DNA sequence from clone RP11-55008 on chromosome 20. Contains
a novel gene encoding a protein kinase, an RPL7 (60S Ribosomal
Protein L7) pseudogene, a CpG island, ESTs, STSs and GSSs, complete
sequence.

ACCESSION
AL359916
VERSION
AL359916.8 GI:9864231
KEYWORDS
HTG; CpG island; protein kinase; RPL7.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 126053)
Griffiths, C.
Direct Submission

COMMENT
Submitted (08-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

On Aug 22, 2000 this sequence version replaced gi:9801431.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone
RP11-55008 it may be shorter because we sequence overlapping

sections only once, except for a 100 base overlap.
The true left end of clone RP11-128M1 is at 125954 in this sequence. The true right end of clone RP4-684024 is at 100 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.
RP11-55008 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6.

FEATURES

Location/Qualifiers

```
1..126053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-55008"
/clone_lib="RPCI-11.2"
```

```
3..106
/note="MIR repeat: matches 51. .154 of consensus"
complement(815..1340)
/note="match: GSS: Em:AQ672921"
971..1046
/note="MER5B repeat: matches 52. .131 of consensus"
1085..1222
/note="L2 repeat: matches 2407. .2551 of consensus"
1341..1847
/note="match: GSS: Em:AQ457593"
1749..1819
/note="L1M1A7 repeat: matches 6214. .6287 of consensus"
1930..2023
/note="MER2 repeat: matches 110. .206 of consensus"
2024..3568
/note="MER5A repeat: matches 1. .1755 of consensus"
3567..3586
/note="MER2 repeat: matches 96. .115 of consensus"
3587..3914
/note="MER2 repeat: matches 2. .344 of consensus"
3915..4010
/note="MER2 repeat: matches 1. .96 of consensus"
4011..4603
/note="L1M1A8 repeat: matches 5632. .6221 of consensus"
complement(4480..4941)
/note="match: GSS: Em:AQ632384"
5309..5412
/note="M1G1G repeat: matches 412. .512 of consensus"
5608..5867
/note="M1G1G repeat: matches 11. .271 of consensus"
6891..7183
/note="AluY repeat: matches 1. .303 of consensus"
7936..8016
/note="MIR repeat: matches 87. .167 of consensus"
8241..8270
/note="15 copies 2 mer ca 100% conserved"
9748..10266
/note="L1M1C4 repeat: matches 7176. .7719 of consensus"
10267..10376
/note="L1M1C4 repeat: matches 7866. .7977 of consensus"
10426..10566
/note="MIR repeat: matches 91. .236 of consensus"
11561..11783
/note="MIR repeat: matches 21. .262 of consensus"
12954..13183
/note="MIR repeat: matches 7. .262 of consensus"
13236..13286
/note="L2 repeat: matches 2445. .2495 of consensus"
13365..13421
/note="L2 repeat: matches 2232. .2288 of consensus"
13443..13675
```

```
/note="MIR repeat: matches 2. .235 of consensus"
13708..13840
/note="L2 repeat: matches 2089. .2228 of consensus"
13958..14021
/note="MIR repeat: matches 80. .146 of consensus"
14640..14874
/note="MIR repeat: matches 2. .262 of consensus"
14966..15086
/note="MERSA repeat: matches 44. .166 of consensus"
15159..15265
/note="MIR repeat: matches 76. .196 of consensus"
15485..15568
/note="MIR repeat: matches 94. .191 of consensus"
15918..16225
/note="AluX repeat: matches 1. .300 of consensus"
16346..16637
/note="AluX repeat: matches 5. .295 of consensus"
16647..16841
/note="L1M1E1 repeat: matches 5474. .5675 of consensus"
16842..17382
/note="LTR36 repeat: matches 1. .612 of consensus"
17387..17423
/note="L1M1E1 repeat: matches 5674. .5710 of consensus"
17806..17944
/note="L1M1E1 repeat: matches 6030. .6167 of consensus"
17979..18090
/note="L2 repeat: matches 1963. .2082 of consensus"
19061..19245
/note="MER3 repeat: matches 2. .184 of consensus"
19288..19814
/note="LTR41 repeat: matches 1. .466 of consensus"
19403..19931
/note="match: GSS: Em:AQ038341"
19858..20168
/note="LTR41 repeat: matches 468. .763 of consensus"
20504..21124
/note="HERVL repeat: matches 433. .1136 of consensus"
21187..21431
/note="HERVL repeat: matches 1155. .1383 of consensus"
21430..21867
/note="HERVL repeat: matches 4883. .5323 of consensus"
21868..22148
/note="AluJb repeat: matches 1. .281 of consensus"
22634..23106
/note="MLT1E repeat: matches 55. .517 of consensus"
23139..23929
/note="L1M1B7 repeat: matches 5273. .6130 of consensus"
23930..24239
/note="AluSq repeat: matches 1. .311 of consensus"
24240..24794
/note="L1M1B7 repeat: matches 4734. .5273 of consensus"
24864..25175
/note="LTR41 repeat: matches 6. .302 of consensus"
25178..25617
/note="LTR41 repeat: matches 237. .663 of consensus"
25810..26105
/note="AluJo repeat: matches 1. .291 of consensus"
26117..26427
/note="AluSq repeat: matches 1. .313 of consensus"
26514..26841
/note="LTR16C repeat: matches 37. .387 of consensus"
26939..27004
/note="MIR repeat: matches 112. .177 of consensus"
27204..27339
/note="MIR repeat: matches 4. .160 of consensus"
28079..28132
/note="MLT1J repeat: matches 103. .156 of consensus"
28133..28315
/note="MER63B repeat: matches 1. .207 of consensus"
28384..28677
/note="LTR16A repeat: matches 122. .418 of consensus"
28734..28853
/note="MIR repeat: matches 68. .191 of consensus"
```

```
repeat_region 29289..29369
/Note="WIR repeat: matches 73..152 of consensus"
repeat_region 29387..29627
/Note="WIR repeat: matches 8..259 of consensus"
repeat_region 29732..30251
/Note="WIR repeat: matches 12..509 of consensus"
repeat_region 30586..30685
/Note="WIR repeat: matches 90..192 of consensus"
repeat_region 30934..31026
/Note="WIR repeat: matches 7864..7949 of consensus"
repeat_region 31072..31167
/Note="48 copies 2 mer ac 69% conserved"
repeat_region 31188..31357
/Note="WIR20 repeat: matches 38..218 of consensus"
repeat_region 32210..32504
/Note="AluY repeat: matches 1..291 of consensus"
repeat_region 32506..32559
/Note="I2 repeat: matches 2645..2697 of consensus"

Query Match 9.1%; Score 272.2; DB 9; Length 126053;
Best Local Similarity 48.0%; Pred. No. 2.4e-61;
Matches 1298; Conservative 0; Mismatches 1353; Indels 54; Gaps 16;

QY 332 TGCCTGAATCCCAACTTTGGAGCGGAGTGGGTGGATCACTTGAAGTCAAGGATT 391
DB 46550 TGCCTGAATCCCAACTTTGGAGCGTGGGTGGCGGATCA--TGAGTCAAGGATT 46607

QY 392 CGAGCAGAGCTGCCACACTGTGAACCTGTCTTACTTAAATAATATAATAGCT 451
DB 46608 CGAGCAGAGCTGCCACACTGTGAACCTGTCTTACTTAAATAATATAATAGCT 46667

QY 452 GGGCATGGTGGTACCTGTAACTCCAGATCTTGGGAGTTGGAGCGGAGGAGTCCG 511
DB 46668 AGGCATGGTGGGTCCTGTAACT--CGACTCTGGGAGCTGGAGCGGAGGAGTCC 46726

QY 512 TTGAACCCGGAGGAGGAGTGGTCAAGTGAACCAAGTGTGCCACTGCATCTCAGTCTGG 571
DB 46727 TTGAACCCGGAGGAGGAGTGGTCAAGTGAACCAAGTGTGCCACTGCATCTCAGTCTGG 46786

QY 572 GCAACAGAGTGAGCTCCATCTCAACCAACCAACCAACCAACCAACCAACCAACCA 630
DB 46787 GTACAGAGGAGTCCATCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 46846

QY 631 -----ATTGAGTGAATGAGTGAAGGAGTGAAGTGTGCAAGTGAAGTGAAGTGA 684
DB 46847 TCATAAATTTGTACAAAGATCCCTTTAAATATAAGACATATAGATTAAGTGA 46906

QY 685 CAGGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 744
DB 46907 AGGGGTAGAGAACATATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 46966

QY 745 AGATGCGGAGGCGCTTCAAGGTGCTGGGATTTGGGAGGATTTGGGAGGATTTGGG 804
DB 46967 TAAATTTCAATAGAGCGGCGGAGTGGCTACGCTGTAAATCCAGCAGTCTTGGGAGG 47026

QY 805 CAGTCTCTGTATTAACAAG--AAAGAGGAGGCGCCATGCTGGTGTGCGCAGGAAGT 861
DB 47027 CTGAGCGGCTGGATTAAGGTCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 47086

QY 862 CAGTAGTAATAGACAGCACCGGTGCTGCTTCCAGCGCACCTAGGCGGAGTGGGAAA 921
DB 47087 CCTGTCTTACTAAGAAATATAAATTTAGTGGCATGTTAGCAGCTGCTGTAGTCCCA 47146

QY 922 CAGACTCACCACAGCTCCAGCCAGAGTGTGTCAGGCGCAGATGGGGAAGCAGCGGA 981
DB 47147 TCTACTCTGGAGCTAGGCA--AGAGAAATCAATTTGAAATCTGGCAGGAGGTTGCA 47204

QY 982 GAAAGTCAAGGTGGGATGGGAGGGGTTCAGGCGAAGAGGGGTTCAGGCGGAGGCTG 1041
DB 47205 GAACTGAGATTGTGCACTGCACTTAGCTGCAAGAGATGAAGTCCCGTCTCATAA 47264

QY 1042 AAGCCCTGGGATCTGTAG--GAATTTAGAGGAGGTACCTGACCGGATGTTTGGTGA 1100
```

```
DB 47265 CAACAACAACAAGCAGAGCAGACTTTCAGAGCAAGAAAAATTTATCAGAAAAAGAGAGC 47324
QY 1101 GATTCAAGGAGTCTTCTCTGGAAGAGAGGCTCTCGAGAGCTGAGACTCATAAGATGAGTGG 1160
DB 47325 ATTACATAACCAATAAATGGTCAATTTCTCCAAAGATACACATAAATATGTCTCCAG 47384
QY 1161 GAGGGTGTTCAGGCGAAGAACCA--GCACCTCAAAAAGCATGCTTTGAGAGAGAGCAT 1219
DB 47385 AAGATATAACAATCTTTAATGTCTATGCACTTAACAACACAGCATCAAAATACATGAGC 47444
QY 1220 CATCAATTAACATGATGAATTTTCAGAGCTGGGCGAGCTGGCTGCTGCTGTAATCCAGC 1279
DB 47445 AAAAACTGACAGAAATGCAAGGAGCGGAGGAGTGGCTCATGCTGCAATCCAGC 47504
QY 1280 ACTTTGAAAGCTGAATGGGAGGATGACTTTGAGCCTAGGCAATTTGACAAAGCTTGGC 1339
DB 47505 ACTTTGGAGGCGCAAGACAGTGGATCACTTGTGCTCCAGGAGTTTACAGCAGGCTGGGA 47564
QY 1340 AACATGGTGAGACCTGCTCCACAAACAAACAAACAA-----CAAAAA 1385
DB 47565 AGTGTAACAATAATCCCATCTCTACAAAAATAAAAAAATTAGCCAGCATGTTGGCTACA 47624
QY 1386 ATCAATTAATCTGTACCATGGGTACAGGTACATAGAAATGACTCAGGCGAGATATGGTG 1445
DB 47625 CTGTGTCTCAGCTCTCAGAGGCTGACGAGGAGGATCACTTGAGCCCCAGAGTTTG 47684
QY 1446 TCCTCTCTACTGTGGAGAGGCGGCTTATATGTCAGTAAAGCAATAGAGGAGGGAAT 1505
DB 47685 AGGCTGAGTGAGTTGTGTTGTGCTGCACTCAGCTCAGCTGAGTGACAGAGAGAGCCCT 47744
QY 1506 ATAAATCTTAATAGAGGTACAGATTTGAGAGCAACACAGGCGACAGGCATATGATG 1565
DB 47745 GTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 47804
QY 1566 AGGCTAAGAGGGAATCAGGGA--AGGCTCTCAGAGAGGTGACATTTAAGCCGGACAT 1624
DB 47805 CTACTATAGTAGAGTCTCAATCTCTCTATCAGCAGTGGCAAAATCAGCAGGCAAA 47864
QY 1625 GAAGGATGAACGAGTTAGTTTCAACAAGGATGGGATGGAAGGGGTGAGAGTGATGGAGC 1684
DB 47865 AGTCAAGTAAAGATATATTTGAACTCAATAGCACCATTATCAACAAGAAATTAATTAACA 47924
QY 1685 AGAGGAATCTGAGGATCATAGGCTTACAGAGGAGTCTTGAAGCCCTTTCAGGAGAGTGA 1744
DB 47925 AATACAGACTACTTCTCAACAACAATAAGATACCCATCTCTCTCAATTTCTGATGAAT 47984
QY 1745 AGAAGACCCAGCGCAGTCTGAGTGGGTTAAGTAAACAAGCTGGAAGCCAGCGGGAATCCCT 1804
DB 47985 ATCACTATGATAGACTACATTTCTGTTCCATTAACAACACCTTAAACACATTTAAAGAATA 48044
QY 1805 G--GTCAATGAGGCGCTGTGAGTCACTGAGTGTGTTGGGCTTTTGTGTTTCTCTGGAG 1862
DB 48045 GAAGTTGGCGCAGCATGTTGGCTCAGCTGTAATCCAGCAGCTTTGTGAGGCGGAGCA 48104
QY 1863 CAGTCAATTTAAGCAGGAAACAGCTGTATTCAGATTTGGGAAGATCCTGTGTTGCTGTC 1922
DB 48105 GCAATATCACTTGAAGTCAAGGATTCAGACAGCAGCTGACCAACATGAGAAACCCCATC 48164
QY 1923 CTGAAGGGGATGAATCTGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGTTGATG 1982
DB 48165 TCTACTAATAATACAAATTTAGCTGGGCATGTTGGGCATACCTGTAAATTTCTAGTACTC 48224
QY 1983 GGGAGCTGGAGGCTCGCGGTGATGGACAGGCTGGGGCCAGGCGGATGGGAGGAGGAGG 2042
DB 48225 GGGAGCTGGAGGAGGAGTCTCTTGAATTTGGGAGG---CAGAGTTGCGGTGAGCTG 48281
QY 2043 AGTAATTTGGAGAGGCTGGGCTCTGGCCGAGGAATGATGTTGGGCTGAAACAGGAG 2102
DB 48282 AGATCAGCAGTGTGATGCCAGCTGGGCAACAGAGCAAAACTCCGTATATAAAAAAANA 48341
QY 2103 AGGAGAGATGCTTAGGCCACTTTGGAAACACAGTAGGCGAGGAGGAGACCCAGGG 2162
DB 48342 AAGAATTAAGAGGAATAGAAATCATACAAATGTCTACTCTTAGGCGCAATGAATTAAC 48401
```

QY 2163 GAAGTGGCCAAAGACACACGACGCTGGCATTTGGACAGGGAAGGTCTGTCTGGAGCAGG 2222
Db 48402 AAGAAATCAATAACAGAAAGACAGCTAGAAATTTTCAAACTATTTTGGAGAGTAACACACA 48461
QY 2223 TGTCTTGGATGAAGGAGGAAATGGTGAGTTCCATCTCTCTCTCTCTCTCTCTCTCAACCTC 2282
Db 48462 CTTTCTGTATAACAGCTCAAGAGAAATCTTTGAGAGAAATTTAACTATTTTGACCTAA 48521
QY 2283 TAAACTACATGGGCGACAGGACCCAGTGGGACTCCATAAATGATGGGATGGGTGGATGGA 2342
Db 48522 GTAA-----AAATGAATAACAAATTTATCAAAATTTGTAGAATGCAGCAAAAG 48570
QY 2343 AGGAAGGAGGAGGAAACAACTCTTCATTCATCTCTGGTTATTTACAGAACAGGCGCAGGTG 2402
Db 48571 CAGTACTTAGAGGGAATTAATAGCATTTGAATATATGTGTAGAGAAAGAGGCTGGGCA 48630
QY 2403 CGGTGCT-CAGCTTGCCATCTAGCACTTTGGGAGGCTGAGTGGGTGGATTACTCTCA- 2460
Db 48631 CGGCGGTACATGCCCTGTAAATCCCAACATTTGGGAGGCCAAGGCGAGTGGATCACTTGAG 48690
QY 2461 GGTCAAGAGTTCAAGACAGGCTAGACAACTAGAGAAACCCCATCTCTTACTGAAGATA- 2519
Db 48691 GGTCAAGAGTTCAAGACAGGCTAGCAGCAATGGCAAAACCTCTCTTACTTAAATAATAC 48750
QY 2520 TAAATAGCTGGGCGTGTAGTGATATGCTGTATCCAGCTAGTGGGAAGCTGAGGC 2579
Db 48751 AAAAATAGCCAGGTGTGTATGSCACACACCTGTAGTTTCACTACTCAAGAGGCTGAGGC 48810
QY 2580 AGGAATTCGTTGAACCCGAGGAGGAGGAGGTTGCGGTGAGCTGAGATCGTCCCATTTGCA 2639
Db 48811 AGTAAATCGCTTCAACCCAGGAGGAGAGATTTGAGTGAGCGGAGATTTGCCACTACA 48870
QY 2640 CTCACGCTGGTGACAAAGCAAGACCTGCTCTCAATATAATAATAATTACAAACAGA 2699
Db 48871 CTCACGCTCGGTGACAGAGCAAGACACTGTCTTAAATAAAGAAAGAAAGAAATA 48930
QY 2700 AGGACCTGGGTATCCAGCTACCTACTTTTTCAGGAGATGTCTCCCTTACCCAGGG 2759
Db 48931 ACTAAATCAATAATCTAAGCTCCATCTTAGGNAATAAATAAATAAAGCAATTAACCC 48990
QY 2760 CAAGGATGGGAGACCACTTTGATTATGATCACTTTATGAGCACCTACTGATGCTCATCC 2819
Db 48991 TAAATGACAGAAATGAAGAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 49050
QY 2820 CTGGGCTAGGCTGG---AATGAGCTCAGATGGAGCCTGAAGAGTCCCTCTCAGGGAACCT 2876
Db 49051 GGGAGTCAATAGGGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 49110
QY 2877 CACTAGAAAGAGGAGAAATCGCGCGGCGGTGCTCAGCGCTGTATCCCAACATT 2936
Db 49111 GGCCAAGGTCAATAAGCCTTGCGCGGCTGTGGTGTGCTCAGCGCTGTATCCCAACATT 49170
QY 2937 TGGAGGCTGAGTGGGTGGATCACAAGGTTCAGAGATCGAGACCATCTCTGGCTAACACA 2996
Db 49171 TGGAGGCGCAGGTGGGCGGATCATGTATTAGAGTTTCAAGACCGCTTGGCCACATG 49230
QY 2997 GTCAA 3001
Db 49231 GTGAA 49235

RESULT 6
AC027272
LOCUS Homo sapiens chromosome 19 clone RP11-486P13, WORKING DRAFT
DEFINITION AC027272
ACCESSION AC027272.4 GI:9838224
VERSION HTG; HTGS PHASE1; HTGS DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AC027272 179607 bp DNA linear HTG 01-SEP-2000
SEQUENCE, 18 unordered pieces.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 179607)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 179607)
Waterston,R.H.
Direct Submission
Submitted (28-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 17, 2000 this sequence version replaced gi:8570236.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0486P13
----- Summary Statistics -----
Sequencing vector: p13; 95%
Chemistry: Dye-primer ET; 95% of reads
Assembly: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 170867 bases at least Q40
Consensus quality: 174124 bases at least Q30
Consensus quality: 176036 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 179350; sum-of-contigs
Quality coverage: 4.55 in Q20 bases; agarose-fp
Quality coverage: 4.80 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1338: contig of 1338 bp in length
* 1339: gap of unknown length
* 1439: contig of 1892 bp in length
* 3331: gap of unknown length
* 3431: contig of 1999 bp in length
* 5429: gap of unknown length
* 5530: contig of 2578 bp in length
* 8108: gap of unknown length
* 8207: contig of 2973 bp in length
* 11181: gap of unknown length
* 11281: contig of 2759 bp in length
* 14040: gap of unknown length
* 17231: contig of 3091 bp in length
* 17331: gap of unknown length
* 17331: contig of 4004 bp in length
* 21335: gap of unknown length
* 21435: contig of 5443 bp in length
* 26877: gap of unknown length
* 26978: contig of 4476 bp in length
* 31553: gap of unknown length
* 31554: contig of 4962 bp in length
* 36516: gap of unknown length
* 36616: contig of 10063 bp in length
* 46679: gap of unknown length
* 46779: contig of 11221 bp in length
* 57999: gap of unknown length
* 58099: contig of 16643 bp in length
* 74742: gap of unknown length
* 74843: contig of 17662 bp in length
* 92504: gap of unknown length
* 92505: contig of 26941 bp in length
* 119545: gap of unknown length
* 119546: gap of unknown length

* 119646 147254: contig of 27609 bp in length
 * 147255 147354: gap of unknown length
 * 147355 179607: contig of 32253 bp in length.

FEATURES

Source
 1..179607
 /organism="Homo sapiens"
 /db_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="RP11-486P13"
 misc_feature
 1..1338
 /note="assembly_name:Contig8"
 1439..3330
 /note="assembly_name:Contig9"
 3431..5429
 /note="assembly_name:Contig10"
 5530..8107
 /note="assembly_name:Contig11"
 8208..11180
 /note="assembly_name:Contig12"
 11281..14039
 /note="assembly_name:Contig13"
 14140..17230
 /note="assembly_name:Contig14"
 17331..21334
 /note="assembly_name:Contig15"
 clone_end:SP6
 vector_side:left
 21435..26877
 /note="assembly_name:Contig16"
 26978..31453
 /note="assembly_name:Contig17"
 31554..36515
 /note="assembly_name:Contig18"
 36616..46678
 /note="assembly_name:Contig19"
 46779..57999
 /note="assembly_name:Contig20"
 58100..74742
 /note="assembly_name:Contig21"
 74843..92504
 /note="assembly_name:Contig22"
 92605..119545
 /note="assembly_name:Contig23"
 119646..147254
 /note="assembly_name:Contig24"
 147355..179607
 /note="assembly_name:Contig25"
 ORIGIN
 Query Match 8.8%; Score 263.8; DB 2; Length 179607;
 Best Local Similarity 67.8%; Pred. No. 4.2e-59;
 Matches 414; Conservative 0; Mismatches 192; Indels 5; Gaps 3;
 QY 2393 AGCCAGGTGGGTCTACGCTTGCATCTAGCACTTTGGAGGCTGAGGTGGGTGA 2452
 DB 25988 AGCCGGGGCGGTCTACGCTGTAATCTTAGCAATTTGGAGGTGGGTGGCGGA 26047
 QY 2453 TTACTCAGGTGAGGTTCAACAGCAGCTAGAACGTAAGAAACCCCATCTCTACT 2512
 DB 26048 TCACCTGAGTCAAGGTTTGAAGACGCTGCGCAACATGCGCAACCCCATCTCTACT 26107
 QY 2513 GAAGATATAAA-ATTAGCTGGGGGTAGTGGCATATGCTGTTGAGTCCAGTCCAGTCCGGA 2571
 DB 26108 AAAAATACAAAGATTAGCCGGGTTGGTGGCATGCTCTGTAATCTCAGCTACTCGGGAG 26167
 QY 2572 GCTGAGGAGGAGATTCGCTTGAACCCGAGAGGAGGTTGCGGTGAGTCCAGTCCGATCGTG 2631
 DB 26168 GCTGAGGAGGAGATTCGCTTGAACCCGAGAGGAGGTTGCAATGAGCCGAGATCATG 26227
 QY 2632 CCATGCACTCAGGCTGGGTGACAAAGCAAGCAAGCTCGTCTCAATA-ATAATAATAA 2690
 DB 26228 CCATGCACTCAGGCTGGGCAACAGGAGGAGCTCTACTCTCAAAATAATAATAATAA 26287

QY 2691 CAAAACAGAGAGCGCTGGGTCACTCCAGCTACCTACTTTTCAGGAGATGACTCCCTT 2750
 DB 26288 AATAAATAATAATAATCATGGAAAGCGAAAAATACATTTATGTGTTTTCCTTCCCTC 26347
 QY 2751 ACCCAAGGGCAAGGATGGGAGAACCAAGTTTGAATATGATTTATTTAGGACCTACTAG 2810
 DB 26348 CCGCAATTAAACCCCTTTTCATTTTCCCATGTCTCTCGCAGTTCCCATTCATCTGCACAA 26407
 QY 2811 TCCTCATCTCCCTGGCTAGGCTGGAAATGGACTCAGATGGAGCTGAAGAGTCCCCCTCAGG 2870
 DB 26408 ATCTTTGATTGCAGAGCGATAAAATAAATGTCATCATGAGCATCCATCGGCAGACGTG 26467
 QY 2871 GAACCTCACTAGAAAGAGGAGGAATCGCGGGCGGGTGGCTCAGCCCTGTATATCCCA 2930
 DB 26468 GCACAAAGTCAATAAAACAAC---TCTAGGCTGGGCACAGTGGCTCACTCTGTATCCCA 26524
 QY 2931 ACCTTTGGGAGGCTGAGGTGGGTGGATCAACAGGTGAGGATCGAGACCATCTGGCT 2990
 DB 26525 GCCTTTAGGAGCGCGAGGAGGTGCATCAAGGTCAAGAGATCAAGACCATCTGGCC 26584
 QY 2991 AACACAGTGAA 3001
 DB 26585 AACATGGTGA 26595

RESULT 7

HS268H5

LOCUS

DEFINITION

Human DNA sequence from clone CTA-268H5 on chromosome 22q13.2-13.3,
 complete sequence.

ACCESSION

AL008718

VERSION

AL008718.23

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

HS268H5

189833 bp

DNA

linear

PRI 04-MAR-2003

Human DNA sequence from clone CTA-268H5 on chromosome 22q13.2-13.3,
 complete sequence.

AL008718

AL008718.23

GI:6006479

HTG.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 189833)

Clark, G.

Direct Submission

Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Oct 1, 1999 this sequence version replaced gi:6002134.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Emi, ENBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 22, constructed by the Sanger Centre Chromosome 22
 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr22	
CTA-268H5 is from the human BAC library described in U-J. Kim et	
al. (1996) Genomics 34, 213-218.	
VECTOR: PBAC108L	
FEATURES	Location/Qualifiers
source	1..188833
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="22"
	/map="q13.2-13.3"
	/clone="CTA-268H5"
	/clone_lib="CIT978SK-A1"
ORIGIN	
Query Match	8.6%; Score 259.2; DB 9; Length 188833;
Best Local Similarity	47.4%; Pred. No. 7,1e-58;
Matches 1296; Conservative	47; Mismatches 1378; Indels 60; Gaps 15;
QY	296 TGATAACAGTGCCTATCAGCCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGG 355
Db	73238 TCAGAAACATGAGCCGAGGCTGGGCACAGTGGCTCACATTTATAATCCAGCACTTTGGG 73297
QY	356 AGGCGGAGGTGGTGGATCATTGAGTCCAGGATTCGAGACCGCTGGCCCAACATGCT 415
Db	73298 AGGCTGAGGCATGTGGATCACCTGAGGTCAGAGGTTCAAGACCGCTGGCCCAACATAGT 73357
QY	416 GAAACCTCTCTACTAAAAATATAAAATTTAGCTGGGCATGTGGTGCCTGACCTGTAA 475
Db	73358 GAAACCCGCTCTACTAAAAATACAAAATTTAGCCAGGTGAGGGTGCACACCTGTAG 73417
QY	476 TCCAGATACCTTGGAGGTGAGGAGGAGAAATCGCTTGAAACCCGGGAGGAGATGTTCG 535
Db	73418 TCCAGCTACTCAGGAGGCTGAGGAGGAGAAATCACTTGAACCTGGGAGGCGAGGTTGC 73477
QY	536 AGTGAACCAAGACTGTGCACCTGCCTCCAGTCCGGAACAGAGTGAAGCTCCATCTCA 595
Db	73478 AGTGAACCGAGTGGCTCCACTGCATCCAGCTGGGCAACAGATGAGATTCGCGCTCA 73537
QY	596 AACAACAAACAAAGCAGTGCCTCATGTAGGATTCAGTGAATTTAGTGAAGCACTGAG 655
Db	73538 AAAAAAATAAAAAATTTAGAAACGTGAGTGGGTGAGTGCAG----- 73582
QY	656 CTTGTGGAAGTGAACACTCAATATCACCAGTTGTAGTATCAGTGATTAACATCAAT 715
Db	73583 ---TTTCTGTGCGCCCGTAACACAGTCACTACAACTGCTGGCTTTAAGCCACAGACAT 73639
QY	716 GATCCAGTAAAGCCCTGAGGGTTTCAGAAAGATGCGCGAGCGCTTTCAAGGTGCTGGGA 775
Db	73640 TTATTTCTCTACAGTTCTAGAACCCAGAGTCAAGAACCAAGGCATCCATGGGGCTGTGC 73699
QY	776 TTGGTGGGCAAGCCCTCGAATAATAGAAACAGTTCTCTGTATTPAACAGAAAGCAGGAG 835
Db	73700 TGCTCCAGAGGCTCCAGGGGAGGAGCGCTCCCTCTTGACAGTCTCTGGTGGCTCCAG 73759
QY	836 GCCCATGCTGGTGTGCTGCCAGGAACCTCAGTAGTAACT-----AAGACAGACCGGTGCTG 890
Db	73760 GCACCTCTTGGCTTGGCGGTGCACCACTCCAAATTTCTGCTCCATCTTTTCAATGGCCT 73819
QY	891 CTTTCCCGCAGCGCACTAGGCGAGTGGGGAACAGACTCACCAACAGTCCCGCCAGAG 950
Db	73820 CTTCTCTGTGTCTACCTGCGACTCAGTCTCCCTGCTGCTTTAAAGTTTCAGGCTTGGCC 73879
QY	951 TGTCTAGGGCCAGATAGGGGAAGCAGCGGAGAAAGGTTCAGGTGGATGGGAGGGGTC 1010
Db	73880 AGGTGCGGTGGCTCACGCTGTAACTCTAAACACTTTTGGGAGACCAAGCGGGTGGATCAC 73939
QY	1011 AGGCGAAGAGGCTCAGGGCCAGGCTGAGGAGAGCCCTGGACTGTAGGAATTTAGAGGA 1070
Db	73940 TTGAGGCGCAGAGTTTGAACACAGCTTGGCCATGATGGTGAACCCCGTCTCTACTTAAA 73999
QY	1071 GGTACCTGCCCGCATGTTTGTGTAGGAGAGATTCAGGAAGTCTTCTCTGGAAGAGGCT 1130

Db	74000 AGACAAAAAATTAGCTGGGTATGTTGGTGCACGCTATAATCCAGCTACTCAGGAGACT 74059
QY	1131 GTCGGAGCTGAGACTATAAGATGAGTGGGAGGGTGTTCAGGAGAGAAAGCAGCAGACC 1190
Db	74060 GAGCAGAGAAATCGCTTGTATCTGGAGCGGAGGTGTGAGTGAGCCAAAGATCGTGCC 74119
QY	1191 TACAAAAGCATGCTTTTGAGAGAAAGCATTCATCCATTCACACTGATGAATTTTCAGACTGG 1250
Db	74120 ACCGACTCCAGCCTGGTGACAGAGTGAGACTCCTTCTCAAAAAAATAAATAATTC 74179
QY	1251 GCACGCT-----GGCTCATCCCTGTGTAATCCAGACACTTTTGGAGAGGTGAATGGG 1300
Db	74180 AGGCTCACACCTGTATAGGCTCACAGCTATAATTCAGCACTTTTGGGAGGCCACGGCAGG 74239
QY	1301 AGGATGACTTGGCCTAGGCATTTGTGACAAAGCCTGGGCAACATGCTCAGACCTGCCTC 1360
Db	74240 CAGATCAGCTGAGGTGAGGAGCCGGCTTGGCCCAACATGTTGAAATGCTGTCTC 74299
QY	1361 CACAAAACAAACAAACAAACAAATCATTTATACCTGGTACCATGGGTACCAG--GTAC 1418
Db	74300 TACTAAAAATACAAA-AATTAGCTGGGCATGGTGCAGTGCCTGTAATCCAGCTGTTC 74358
QY	1419 ATAGAAATGACTCAGCGAGATATGGTGTCTCTCTACTGTGGGAGAGCGGCTTATAC 1478
Db	74359 GGGAGGCTGAGCGAGGAGAAATCGTTTGAACCCAGGAGCGGGGTTTACAGTGAGCCGAGA 74418
QY	1479 TGCAGTAAGACNAATAGAGGGGGAATATAATCTTAAATGAGAGTACAGATTTTGAGAG 1538
Db	74419 TTGCCCCACCCGACCTCCAGCCTGGCGCAAAAGCAAGACTGTCCCTCCCAAAAAAANA 74478
QY	1539 CAAACACAGGGCAGCAGGCATATGTACGAGGGTAAAGAGGGAATCAGGGAAGCTTCTCAG 1598
Db	74479 AAGTTCACTTGCATTTGAAGTGAACCCGCTGTAATCCAGGATGCTGTCA--CCTCAAGA 74536
QY	1599 AGAAGGTGACATTTAAGCCGGACATGAAGAGATGAAGAGTTAGTTTACCAGGATGCGA 1658
Db	74537 GACTGAGCTTAATCACTCTGCAAGAATTTTCTTTCCTCAATAGGTCCTCGCATTCACA 74596
QY	1659 TGGAAAGGGTGAAGATGATGAGCGCAGAGGGAATCGCAGGATCATAGCCCTAGACAGGG 1718
Db	74597 GGTAGACATCTCATTTTGAGGAAGTATGGGACA----AATTCAACTCACTGCTGCAGC 74652
QY	1719 GATCTGTAGCCCTTGAGGAAGTGAAGAGAACAGCGCAGTCGTAGTGGTGTAAAGTAAC 1778
Db	74653 CTTCTCCACCACTTTTGTGTAGTACAGCATCTGTCTGTCTCTCTCATGATGGGTTTT 74712
QY	1779 AAGCTGAGAACCGAGGGAATCCCTGGTTCATGAGGCGCTGTGAGTCACTGTCAGAGTGT 1838
Db	74713 TCCTCTGGGACCCACCTTGTGCCCATTTTCAGTTTACAGAGCTGAGCCAGGCTTGCCCA 74772
QY	1839 TTGGGCTTTTGTTTTTCCTGGGAGCAGTCGATTTTAAAGCAGGGAACAGCTGTATTTCAGAG 1898
Db	74773 TTAATCAGAGCATCTCCCTCGCCATAGTAGTACAGGATGGTCCAGTCAGGSCAA 74832
QY	1899 TTGGGAAGATCCTGTGGTGTCTGCTTGAAGGGGATGAACACTGAGAGCTGAGGCCACAGGG 1958
Db	74833 TCCAGGACTCTGATGTCTTTGACGTAAGGCCACTTAAGAGGACAGCTGCGCAGCC--GCC 74890
QY	1959 TGAATGGGAGGATCCAGGCTGATGGGAGGCTGGAGGTCCCGGCTGATGGACACAGGCT 2018
Db	74891 ATCTTTCCAGACCCATCTTGTAGGCGACCTGGCTGAGAAATGGAGTCCACACAGAG-- 74948
QY	2019 GGGCCAGGGATGGGAGGAGAGTAAATTTGGGAGAGGCTGGGGCTCTGGCCGAGGAA 2078
Db	74949 GAGAACGGGCCAGAGATGGAAGTGAAGCCAGCTTTTCAGTTCCGGAGCCAATACAC 75008
QY	2079 TGGATGGTGGGTGAACACAGGGAGAGAGATGCTTTAGGCCACTTTTGGNACACAGTAGG 2138
Db	75009 CCCCTTTCTGCTGAAGCCAGCTGAGAGGGGTTTCTGGCCCTTACTCATTTGTAGGTTTT 75068
QY	2139 GGAAGACAGGACACCCNAGGGAGTGCCTCAAGAGACACAGCAGGCTGGCATTTGA 2198
Db	75069 GAAGGAAGCAAGGTTTAAAGGAAAAGACACACACACACACACACACACACACAC 75128

Mon Sep 20 09:09:29 2004

```

repeat_region      /rpt_family="AluSg"
11823..11923
repeat_region      /rpt_family="MLT1K"
complement(11936..12253)
repeat_region      /rpt_family="AluJb"
complement(12268..12576)
repeat_region      /rpt_family="AluSc"
complement(12610..12817)
repeat_region      /rpt_family="MLT1K"
complement(12820..12988)
repeat_region      /rpt_family="AluSp"
complement(12990..13283)
repeat_region      /rpt_family="AluSg"
complement(13287..13418)
repeat_region      /rpt_family="MLT1K"
13488..13755
repeat_region      /rpt_family="AluJb"
13761..13790
repeat_region      /rpt_family="CA)n"
13794..13922
repeat_region      /rpt_family="AluSg/x"
13924..14060
repeat_region      /rpt_family="AluSg"
14061..14370
repeat_region      /rpt_family="AluSx"
14371..14534
repeat_region      /rpt_family="AluSg"
14541..14717
repeat_region      /rpt_family="MIR"
complement(15059..15160)
repeat_region      /rpt_family="MIR"
15649..15720
repeat_region      /rpt_family="MIR"
16036..16085
repeat_region      /rpt_family="MIR"
16109..16410
repeat_region      /rpt_family="AluSg"
complement(16447..16654)
repeat_region      /rpt_family="MIR3"
17064..17245
repeat_region      /rpt_family="Li"
complement(17291..17399)
repeat_region      /rpt_family="MIR3"
complement(17799..17909)
repeat_region      /rpt_family="LIME3A"
18318..18373
repeat_region      /rpt_family="GC rich"
complement(18508..18633)
repeat_region      /rpt_family="FLAM_A"
complement(18636..18797)
repeat_region      /rpt_family="LIME3A"
complement(18798..18869)
repeat_region      /rpt_family="MLT1B"
18872..19003
repeat_region      /rpt_family="AluJo/FLAM"
19005..19294
repeat_region      /rpt_family="AluSx"
complement(19586..19802)
repeat_region      /rpt_family="U3"
20084..20410
repeat_region      /rpt_family="AluJo"
complement(20415..20784)
repeat_region      /rpt_family="LIME4"
complement(20785..21046)
repeat_region      /rpt_family="AluJb"
complement(21048..21102)
repeat_region      /rpt_family="LIME4"
21103..21398
repeat_region      /rpt_family="AluSg"
complement(21399..21594)
repeat_region      /rpt_family="LIME4"
21595..21903
repeat_region      /rpt_family="AluSg"

```

```

repeat_region      complement(21904..22221)

Query Match      8.6%; Score 258; DB 9; Length 40948;
Best Local Similarity 65.1%; Pred. No. 1.5e-57;
Matches 396; Conservative 0; Mismatches 210; Indels 2; Gaps 1;

QY 2394 GGCACGGTGGGTGCTCACGCTTGCCTTCCCATTTCTAGCACCTTTGGGAGGCTGAGTGGGTGGAT 2453
DB 25635 GCCAGGCTTGGTGGCTCACGCTTGTATCCAGCACCTTTGGGAGGCCAAGCGCGCGGAT 25576
QY 2454 TACCTCAGGTGAGGATTCAAGACCCAGCTAGACCAAGTAGAGAAACCCCATCTCTACTG 2513
DB 25575 CACCTGAGGTCGGGAGTTGGAGACCGCTCACCACATGGAGAAACCCCATCTCTACTA 25516
QY 2514 AAGTATATAAATTAGCTGGCGTAGTGCATATCCCTGTATCCAGCTAGTCCGGAGGC 2573
DB 25515 AAAATACAAAATTAGCCAGGCATGGTGCAATGCCCTGTATCCAGCTATCTCGGAGGC 25456
QY 2574 TGAGGCAGGAGATCGCTTTGAACCCGAGAGCAGAGTTGCGGTGAGCTGAGATCGTGCC 2633
DB 25455 TGAGGCAGGAGATTGCTTTGAACCCGAGAGTTGCGGTGAGCTGAGATCATGCC 25396
QY 2634 ATTCACTCCAGCTGGGTGACAAAGCAGACCTCGTCTCAATATATATATATATACAA 2693
DB 25395 ATTCACTCCAGCTGGAGAGAGTGAACCTCCGTCTCAAAAATAAAAATATATAT 25336
QY 2694 AACGAGGAGGCTGGGTGATCCAGCTACCTACTTTTTCAGGAGATGCTACTCCCTTACC 2753
DB 25335 TTGTGTTAGCATGGTGACTCAGCTTACCTGTATCTAGACTT--TGGGAGGCC 25278
QY 2754 CAAGGCAAGGATGGGAGAACCCAGTTTGTATATGTCATTTATTTAGGACCTTCTAGTCC 2813
DB 25277 GAGGTAGGAGATTGCTTTGAGGCCAGGAGTTTCAAGACTAGCTGGCGCAACATAGCAAGAT 25218
QY 2814 TCATCCCTGGGTAGGCTGGATGAGCTCAGATGGAGCTGAGAGTCCCTCCCTCAGGAA 2873
DB 25217 CCCATCTCTACAGAAAGTAAATTTTAAATAATATGACATTTTGGCTGCTGATGATA 25158
QY 2874 CTTCACTAGAAAGAGGAGGATCGCGGGCGGTGGCTCACGCCCTGTATATCCCAACA 2933
DB 25157 AAAATTTAAAGAAAGAAAGAGGCGCGGCGGTGGCTCACGCCCTGTATATCCAGCA 25098
QY 2934 CTTTGGGAGGCTGAGTGGGTGGATCAAGGTGAGGATGAGACATCTCGCTAAC 2993
DB 25097 CTTTGGGAGGCTGAGTGGGTGGATCATGAGGTGAGGATCAAGACCATCTCGCTAAC 25038
QY 2994 ACAGTGAA 3001
DB 25037 GCGGTGAA 25030

```

```

RESULT 10
AC144507
LOCUS
DEFINITION
AC144507
AC144507.1 GI:30103061
VERSION
HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 41867)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,

```

Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (25-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L29228

Center clone name: 80066_E_9

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 35848: contig of 35848 bp in length

* 35949 gap of 100 bp

* 35949 41867: contig of 5919 bp in length.

Location/Qualifiers

1. .41867

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosomes="17"

/map="17"

/clone="XXfos-80066E9"

/clone_lib="Human Female Fosmid Library"

ORIGIN

Query Match

Best Local Similarity 8.6%; Score 258; DB 2; Length 41867;

Matches 396; Conservative 0; Mismatches 210; Indels 2; Gaps 1;

```

QY 2394 GGCGAGGTGGTGTGCTGAGTTCAGACACCGCTAGACACGCTAGAGAACCCCATCTCTACTG 1;
Db 5394 GGCGAGGTGGTGTGCTGAGTTCAGACACCGCTAGACACGCTAGAGAACCCCATCTCTACTG 2453
QY 2454 TACCTCAGTTCAGTTCAGTTCAGACACCGCTAGACACGCTAGAGAACCCCATCTCTACTG 2513
Db 5454 CACCTGAGTTCGGAGTTCGGACACCGCTTCACACATGGAGAACCCCATCTCTACTA 5513
QY 2514 AGATATATAAATAGTGGCGGTAGTGGCATATGCTGTAATCCAGCTAGTCGGGAAGC 2573
Db 5514 AAAATACAAAATAGCCAGCATGTGGCAGCATGCTGTAATCCAGCTACTCGGAGGC 5573
QY 2574 TGAGCAGCAGGAATCGCTTGACCCGAGGAGGAGGTTGCGGTGAGCTGAGATCGTGCC 2633
Db 5574 TGAGCAGCAGGAATTCCTTGAACCCGGAGGTGGAGGTTGCGGTGAGCTGAGATCATGCC 5633

```

FEATURES

source

```

QY 2634 ATTGCACCTCCAGCCTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATTACAA 2693
Db 5634 ATTACACTCCAGCCTGGGTGACAAAGCAAGACCTCGTCTCAATAATAATAATAATTAT 5693
QY 2694 AACAGAAGGAGCCTGGGTGATCCCGAGCTACCTCTTTTCAGGAGATGCTACTCCCTTACC 2753
Db 5694 TTGTGGTTAGGCATGGTGCATCAGCTTACACCTGTGAATCTCCTAGAACTT--TGGAGGGCC 5751
QY 2754 CAAGGGCAAGGATGGGAGAACCAAGTTTGTATATGCAATTTATTGAGCACCTACTGATGCC 2813
Db 5752 GAGGTAGGAGGATTTGCTTGGCCAGGAGTTCAAGACTAGCTGGGCAACATAGCAAGAT 5811
QY 2814 TCATCCCTGGGTAGGCTGGAAATGGAATCAGATGGAGCTTCAAGAGTCCCTCAGGAA 2873
Db 5812 CCCATCTCTACAGAAAAGTAAATTTTAAATAATATGACATTTTGAGCTGCTGATGAAA 5871
QY 2874 CCTCACTAGAAAAGAGGAGGAATCGGCCGGCGGTGGCTCAGCGCTGTAAATCCGACAA 2933
Db 5872 AATATTTAAAGAAAGAAAGAGAGGGCCGGCGGTGGCTCAGCGCTGTAAATCCGACAA 5931
QY 2934 CTTTGGGAGCTGAGTGGGTGGATCAAGGTCAAGGTCAAGGATCGAGACCATCTCGCTAAC 2993
Db 5932 CTTTGGGAGCTGAGTGGGTGGATCATGAGTCAAGGATCAAGACCATCTCGCTAAC 5991
QY 2994 ACAGTGAA 3001
Db 5992 GCGGTGAA 5999

RESULT 11
HSAJ9617/c
LOCUS
DEFINITION
Homo sapiens chromosome 17 sequence from PAC RP1-149D14 and cosmid
40M01, complete sequence.
ACCESSION
AJ009617 AL035361
VERSION
AJ009617.5 GI:29466470
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Sudbrak, R., Mueller, I., Kosiura, A., Radelof, U., Hennig, S.,
Ramser, J., Francis, F., Steffens, C., Klein, M., Seranski, P.,
Poustka, A., Lehrach, H. and Reinhardt, R.
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 153609)
AUTHORS
MPIMG.
TITLE
Direct Submission
JOURNAL
Submitted (17-JUL-1998) MPIMG, Abt. Lehrach, Max Planck Institut
fuer Molekulare Genetik, Ihnestrasse 73, Berlin, 14195, Germany
COMMENT
On Apr 1, 2003 this sequence version replaced gi:9211531.
contig 01 1. .153609
Clone received from the Resource Centre of the Human Genome Project
at the Max-Planck-Institute for Molecular Genetics.
FEATURES
Location/Qualifiers
1. .153609
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p11.2"
/clone="COSMID 40M01"
institute: Roswell Park Cancer Institute, creator: Pieter
de Jong, P. Ioannou 40M01: L4/FS17 (human chromosome 17
specific cosmid library)
/note="region between markers D17S842-D17S953"
misc_feature
1
/note="Sp6_end:RP1-149D14"
153609
misc_feature
153609
/note="T7_end:COSMID 40M01"
ORIGIN

```

Query Match 8.6%; Score 258; DB 9; Length 153609;
Best Local Similarity 65.1%; Pred. No. 1.5e-57;
Matches 396; Conservative 0; Mismatches 210; Indels 2; Gaps 1;

QY 2394 GCCAGGTGGGCTGCTCAGCTTCCATCTAGCACTTGGAGGCTGAGTGGGTGGAT 2453
DB 15817 GCCAGGCTGGTGGCTCAGCTTCCATCTAGCACTTGGAGGCTGAGTGGGTGGAT 15758

QY 2454 TACCTCAGGTTCAGAGTTCAGACAGCTAGACAAAGCTAGAGAAACCCCATCTCTACTG 2513
DB 15757 CACCTGAGTTCGGAGTTCAGACAGCTCACCACATGAGAAACCCCATCTCTACTA 15698

QY 2514 AAGATATAAATAGCTGGGCTAGTGCATATGCTGTGTAATCCAGCTAGTCCGGAAGC 2573
DB 15697 AAAATACAAATAGCCAGGCTGTCGACATGCTGTGTAATCCAGCTAGTCCGGAAGC 15638

QY 2574 TGAGGAGGAGAAATCGCTTGAACCGGAGGAGGAGGTTGGCTGAGCTGAGATCGTGCC 2633
DB 15637 TGAGGAGGAGAAATCGCTTGAACCGGAGGAGGAGGTTGGCTGAGCTGAGATCGTGCC 15578

QY 2634 ATTGCATCTCCAGCTTGGGCTGACAAAGCAAGCACTGCTCTCAATATAATAATATACAA 2693
DB 15577 ATTACATCTCCAGCTTGGGCTGACAAAGCAAGCACTGCTCTCAATATAATAATATAT 15518

QY 2694 AACAGAGAGGAGCTGGGCTATCCAGCTACCTATCTTTTCAGAGAGATGATCTCCCTACC 2753
DB 15517 TTGTGTTAGGATGGTCACTCAGCTTACACCTGTGTAATCTCTAGAACTT--TGGAGAGCC 15460

QY 2754 CAAGGCAAGAGTGGGAGACCACTTTGATATGCTATTTATGAGCACTTACTGAGTCC 2813
DB 15459 GAGGTAGGAGGATCTTGGAGCCAGGAGTTCAAGACTAGCTTGGGCAACATAGCAAGAT 15400

QY 2814 TCATCCCTGGCTAGGCTGGAATGCACTCAGATGAGGCTGCAAGAGTCCCTCTCAGGAA 2873
DB 15399 CCCATCTCTACAGAAAGTAATTTTAAATATATGACATTTTGGAGTGCATGATGAA 15340

QY 2874 CCTCACTAGAAAGAGGAGGAATCGGCGGGCGGGTGGCTCAGCCCTGTATCCCAACA 2933
DB 15339 AAAATTTAAAGAAAGAAAGAGGCGGGCGGGTGGCTCAGCCCTGTATCCCAACA 15280

QY 2934 CTTTGGGAGGCTGAGTGGTGGATCAACAGGTCAGGAGATCGAGCACTCTCGCTAAC 2993
DB 15279 CTTTGGGAGGCTGAGTGGGCGGATCATGAGTTCAGGATCAAGACCATCTCGCTAAC 15220

QY 2994 ACAGTCAA 3001
DB 15219 GCGGTGAA 15212

RESULT 12
AC007952 AC007952 158103 bp DNA linear PRI 03-APR-2003
LOCUS Homo sapiens chromosome 17, clone RP11-160E2, complete sequence.
DEFINITION AC007952
ACCESSION AC007952
VERSION AC007952.15 GI:29501895
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 158103)
TITLE Homo sapiens chromosome 17, clone RP11-160E2
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 158103)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A.,
Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
Cook, P., DeArelano, K., Depayre, E., Devon, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,
Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
Hagos, B., Heaford, A., Horton, L., Howland, J.C., Jones, C., Kann, L.,

TITLE
JOURNAL
REFERENCE
AUTHORS

Karatas, A., Lebozky, J., Lieu, C., Locke, K., Macdonald, P.,
Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

Direct Submission
Submitted (02-JUL-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 158103)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gordon, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J.,
Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (20-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 158103)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Jones, C.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (03-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 3, 2003 this sequence version replaced gi:25140154.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submission@genome.wi.mit.edu
----- Project Information
Center project name: L913
Center clone name: 160_E_2

This clone contains two copies of a large inverted duplication from
about base pair 55,000 to 82,000 and about base pair 130,000 to

158,000. Bases 154,031 to 154,092 are represented here by the corresponding bases of the first copy of the duplication.

```
FEATURES
  source
    1. .158103
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="17"
    /map="17"
    /clone="RP11-160E2"
    /clone_lib="RPC1-11 Human Male BAC"
    331. .374
    /rpt_family="FLAM A"
    16810. .17108
    /rpt_family="AluY"
    complement(17112. .17416)
    /rpt_family="AluSc"
    17717. .18062
    /rpt_family="MLT1A0"
    complement(18133. .18175)
    /rpt_family="MIR3"
    18511. .18763
    /rpt_family="L2"
    complement(19346. .19460)
    /rpt_family="MIR3"
    19519. .19672
    /rpt_family="MIR"
    19708. .19847
    /rpt_family="L2"
    complement(20576. .20663)
    /rpt_family="MIR"
    20991. .21153
    /rpt_family="L2"
    complement(21693. .21794)
    /rpt_family="MIR3"
    22878. .22898
    /rpt_family="GC rich"
    complement(22968. .23085)
    /rpt_family="MIR"
    23582. .23882
    /rpt_family="AluSx"
    complement(24182. .24427)
    /rpt_family="AluSc"
    24646. .25088
    complement(24646. .25088)

repeat_region 16677. .16802
repeat_region 16810. .17108
repeat_region 17717. .18062
repeat_region 18511. .18763
repeat_region 19519. .19672
repeat_region 19708. .19847
repeat_region 20991. .21153
repeat_region 22878. .22898
repeat_region 23582. .23882
repeat_region 24646. .25088

Query Match 8.6%; Score 258; DB 9; Length 158103;
Best Local Similarity 65.1%; Pred. No. 15e-57;
Matches 396; Conservative 0; Mismatches 210; Indels 2; Gaps 1;

QY 2394 GCCAGGTCGGTGCTCAGCTTGCCATTCAGCACTTGGGAGGCTGAGTGGGTGGAT 2453
Db 36314 GCCAGGCTTGGTGCTCAGCTTGCCATTCAGCACTTGGGAGGCTGAGTGGGTGGAT 36373

QY 2454 TACCTCAGTCAGGAGTTCAAGACAGCCTAGACAAAGTAGAGAAACCCCATCTCTACTG 2513
Db 36374 CACCTGAGTCGGGAGTTGGAGACCCAGCCTCACCACAGAGAAACCCCATCTCTACTA 36433

QY 2514 AAGATATAAATTTAGCTGGCGTAGTGCCATATGCTGTAATCCAGCTAGTCCGGAAGC 2573
Db 36434 AAAATACAAATTTAGCCAGGCATGTTGGCAGATGCTGTAATCCAGCTAGTCCGGAAGC 36493

QY 2574 TGAGGAGGAGAAATCGTTGAACCCGAGAGGAGGTTGCGGTGAGCTGAGATCGTGCC 2633
Db 36494 TGAGGAGGAGAAATCGTTGAACCCGAGAGGAGGTTGCGGTGAGCTGAGATCGTGCC 36553

QY 2634 ATTGCACTCCAGCTCGGTGCAAAAGACAGACCTGCTCAATATAATAATAATAATAA 2693
Db 36554 ATTCACTCCAGCTCGGTGCAAAAGACAGACCTGCTCAATATAATAATAATAATAA 36613

QY 2694 AACAGAAGGAGCTGGGTGATCCAGCTACCTACTTTTTCAGGAGAAATGCTACCTTACC 2753
Db 36614 TTGTGTTAGGATGCTGACTCAGCTTACCTGTAATCTTAGAATCTT-TGGGAGGCTC 36671

QY 2754 CAAGGGCAAGATGGGAGAACCAAGTTTGAATATGCAATTTATTTAGACACCTACTGAGTCC 2813
Db 36672 GAGGTAGGAGGATGCTTTGAGGCCAGGAGTTTCAAGACTAGCTGGGCAACATAGCAAGAT 36731

QY 2814 TCATCCCTGGGCTAGGCTGGATGCACTCAGATGAGGCTGAAGAGTCCCTCTAGGAA 2873
Db 36732 CCCATCTCTACAGAAAGTAAATTTTAAATAATATGACATTTTGGAGCTGATGATGAA 36791

QY 2874 CCTCACTAGAAAGAGGAGGAAATCGCGCGGCGGCTGCTCAGCCTGTATCTCCACA 2933
Db 36792 AAATTTTAAAGAAAGAAAGAGGCGCGGCGGCTGCTCAGCCTGTATCTCCACA 36851
```


<p>TITLE JOURNAL REFERENCE AUTHORS</p>	<p>Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tessier, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.</p> <p>Direct Submission Submitted (21-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 3 (bases 1 to 176584)</p>	<p>4754 4854 6191 6192 6292 7659 7659 9247 9248 9348 10792 10891 10892 13201 13202 13302 15490 15491 17422 17423 17523 19982 20082 21487 21587 21588 23067 23166 25655 25656 25755 27997 28097 30125 30225 30226 33061 33160 37147 37247 37248 40807 40907 40908 43946 44045 46338 46339 46439 50865 50866 50965 50966 56185 56286 56286 61204 61304 61305 66925 67024 70900 70901 71000 76151 76251 76252 81819 81820 81919 87341 87342 87442 92459 92559 98672 98672 98772 98773 103721 103722 103822 112632 112632 121203 121203 121303 121303 127799 127799 140599 140599 140600 140700 156653 156753</p>	<p>4853: gap of 100 bp 6191: contig of 1338 bp in length 6291: gap of 100 bp 7658: contig of 1367 bp in length 7758: gap of 100 bp 9247: contig of 1489 bp in length 9248: gap of 100 bp 10791: contig of 1444 bp in length 10891: gap of 100 bp 10892: contig of 2310 bp in length 13201: gap of 100 bp 13302: contig of 2189 bp in length 15490: gap of 100 bp 15491: contig of 1832 bp in length 17422: gap of 100 bp 17423: contig of 2459 bp in length 17523: gap of 100 bp 19982: gap of 100 bp 20082: contig of 1406 bp in length 21487: gap of 100 bp 21587: contig of 1479 bp in length 21588: gap of 100 bp 23067: contig of 2489 bp in length 23166: gap of 100 bp 25655: gap of 100 bp 25656: contig of 2241 bp in length 25755: gap of 100 bp 27997: gap of 100 bp 28097: contig of 2029 bp in length 30125: gap of 100 bp 30225: contig of 2835 bp in length 30226: gap of 100 bp 33061: contig of 3987 bp in length 33160: gap of 100 bp 37147: contig of 100 bp 37247: gap of 100 bp 37248: contig of 3560 bp in length 40807: gap of 100 bp 40907: contig of 3038 bp in length 40908: gap of 100 bp 43946: gap of 100 bp 44045: contig of 2293 bp in length 46338: gap of 100 bp 46339: contig of 4427 bp in length 50865: gap of 100 bp 50866: contig of 5220 bp in length 50965: gap of 100 bp 50966: contig of 4919 bp in length 56185: gap of 100 bp 56286: contig of 100 bp 61204: gap of 100 bp 61304: contig of 5620 bp in length 61305: gap of 100 bp 66925: contig of 3876 bp in length 70900: gap of 100 bp 70901: contig of 5151 bp in length 71000: gap of 100 bp 76151: contig of 5568 bp in length 76251: gap of 100 bp 76252: contig of 5422 bp in length 81819: gap of 100 bp 81820: contig of 5017 bp in length 81919: gap of 100 bp 87341: contig of 6114 bp in length 87342: gap of 100 bp 87442: contig of 4949 bp in length 92459: gap of 100 bp 92559: contig of 8811 bp in length 98672: gap of 100 bp 98672: contig of 100 bp 98772: gap of 100 bp 98773: contig of 8471 bp in length 103721: gap of 100 bp 103722: contig of 6396 bp in length 103822: gap of 100 bp 112632: contig of 12800 bp in length 112632: gap of 100 bp 121203: contig of 15953 bp in length 121203: gap of 100 bp 121303: contig of 15953 bp in length 121303: gap of 100 bp 127799: gap of 100 bp 127799: contig of 15953 bp in length 140599: gap of 100 bp 140599: contig of 15953 bp in length 140600: gap of 100 bp 140700: contig of 15953 bp in length 156653: gap of 100 bp 156753: contig of 19832 bp in length.</p>	<p>Location/Qualifiers 1. .176584 /organism="Homo sapiens"</p>
---	--	--	--	--

<p>TITLE JOURNAL COMMENT</p>	<p>Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www.seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information ----- Center project name: L9984 Center clone name: 297_C_16 ----- Summary Statistics ----- Sequencing vector: M13; M77815; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Assembly program: Phrap; version 0.960731 Consensus quality: 159055 bases at least Q40 Consensus quality: 167576 bases at least Q30 Consensus quality: 170993 bases at least Q20 Insert size: 165000; agarose-fp Insert size: 172884; sum-of-contigs Quality coverage: 4.1 in Q20 bases; agarose-fp Quality coverage: 3.9 in Q20 bases; sum-of-contigs</p>	<p>1405: contig of 1405 bp in length 1406: gap of 100 bp 1506: contig of 1657 bp in length 3162: gap of 100 bp 3262: gap of 100 bp 4753: contig of 1491 bp in length</p>
---	---	---

NOTE: This is a 'working draft' sequence. It currently consists of 38 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

		/mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="22" /map="22" /clone="RP11-297C16" /clone_lib="RPC1-11 Human Male BAC" 1.1405 /note="assembly_fragment" 1506.3162 /note="assembly_fragment" 3263.4753 /note="assembly_fragment" 4854.6191 /note="assembly_fragment" 6292.7658 /note="assembly_fragment" 7159.9247 /note="assembly_fragment" 9348.10791 /note="assembly_fragment" 10892.13201 /note="assembly_fragment" 13302.15490 /note="assembly_fragment" 15591.17422 /note="assembly_fragment"																																																																																																																															
		Query Match Best Local Similarity 47.4%; Pred.No.1.5e-57; Matches 1295; Conservative 0; Mismatches 1380; Indels 59; Gaps 15;																																																																																																																															
QY	296	TGATAACAGTGGCCATCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGG	355	128084	TCAGAAACATGAGCCAGGCTGGGCACAGTGGCTCACATTATATCCAGCACTTTGGG																																																																																																																												
Db	128084	TCAGAAACATGAGCCAGGCTGGGCACAGTGGCTCACATTATATCCAGCACTTTGGG	128143	QY	356	AGCGGAGGTGGGTGATCATCTTGAAGTTCAGAGTTCGAGACGAGCTGGCCAAACATGGT	415	128144	AGCGTGAGGCATGTGGATCACCTGAGTCAAGTTCAGGAGTTCAGACCCAGCTGGCCAAACATAGT	128203	QY	416	GAACCCCTCTCTACTATAAAATATAAAATATAGCTGGGCTGAGTGGGCTGACCTGTAA	475	128204	GAACCCCTCTCTACTATAAAATATAAAATATAGCTGGGCTGAGTGGGCTGACCTGTAA	128263	QY	476	TCCAGATCTTTGGGAGTTGAGGAGGAGATCGCTTGAACCCGGGAGGAGCATGTTC	535	128264	TCCAGTACTCAGGAGGCTGAGGAGGAGATCACTTGAACCTGGAGGCGGAGTTGC	128323	QY	536	AGTGACCAAGACTGTCACCTGCACTCCAGTTCGGGCAACAGAGTGAGCCTCCATCTCA	595	128324	AGTGACCAAGACTGTCACCTGCACTCCAGTTCGGGCAACAGAGTGAGCCTCCATCTCA	128383	QY	596	AACAAACAAACAAACAGCAGTCCCATCATGTAGGATTCAGTTCAGTTCAGTTCAGTTCAG	655	128384	AACAAACAAACAAACAGCAGTCCCATCATGTAGGATTCAGTTCAGTTCAGTTCAGTTCAG	128428	QY	656	CTTTGTGCAAGTGCATCTACTAATCAACAGGTTGTAGTATCAGTGTATACCACTCAAT	715	128429	CTTTGTGCAAGTGCATCTACTAATCAACAGGTTGTAGTATCAGTGTATACCACTCAAT	128485	QY	716	GATCAGGTAAAGCCCTGAGGGTTGAGAAAGTGCAGGAGCCCTTCAAGGTGCTGGGA	775	128486	TTATTTCTCAGAGTTCAGAGGAGGAGGAGTTCAGAAACCAAGGATCCATGGGGTGTGC	128545	QY	776	TTGTGGGCAAGCCCTCGAATATAGAACAGTTCCTGTATTAACAGAAAGCAGGAG	835	128546	TGCTCCAGAGCTCAGGGGAGGAGGAGTTCCTGCTCTTCAGGTTCTTGTGGGTCCAG	128605	QY	836	GCCCATGCTGGTGTGTCAGGAACTCAGTAGTAATCTCTCTGATTAACAGAAAGCAGGAG	890	128606	GCATCTCTGGCTTGGCTGACCACTCCCAATTTCTGCTCCATCTCTTCATGAGGCT	128665	QY	891	CTTCCCAAGCGCACTAGGCCAGTGGGAAACAGACTCACCACAGTCCCGCCAGAG	950	128666	CTTCCCAAGCGCACTAGGCCAGTGGGAAACAGACTCACCACAGTCCCGCCAGAG	128725																																																								
Db	128666	CCTCCTCTGTCTACTGGACTCAGTCTCCCTCTGCTCTTAAAGTTTCAGGCTTGCC	128725	QY	951	TGTCAGGGCCAGATGGGAGACACGGGAGAAAGTTCAGGTTGGGAGGGGCTC	1010	128726	AGGTGCGGTGGCTCAGCGCTGTAATCTTAAACACTTTGGAGACCAACGGCGGTGATCAC	128785	QY	1011	AGGCAAGAGGGGTGAGGGCCAGGCTGAGGGAAGCCCTGGAGCTGTAGAAATTTAGAGA	1070	128786	TTGAGGCGAGGAGTTTGAACACAGCTTGGCCATGATGGTGAACCCCTCTCTACTAAA	128845	QY	1071	GGTACCTGACCCCGCATGTTTGGTGAGGAGATTCAGGAAGTTCCTTGGGAAGAGGCT	1130	128846	AGACAAAATTTAGCTGGTATGTTGTGACGCTTATATCCAGCTACTCAGGAGCT	128905	QY	1131	GTGAGACTGAGACTCATAGATGAGTGGGAGGTTTCCAGGACGAAAGACCCAGCACC	1190	128906	GAGCAGAGAAATCGCTTGTATCCTGGAGGCGAGGTTCAGTGAGCCAGATCTGTC	128965	QY	1191	TACAAAGCATGACTTTGAGAGAAGCATTCATCAATTCATGATGATTTTCAGATGG	1250	128966	ACCGACTCCAGGCTGGGTGACAGAGTGAGACTCTTCTCAAAAAAATAAATAATTC	129025	QY	1251	GCACGCT-----GGCTCATGCTGTAATCCAGCACTTTTGGAGGCTGAATGGG	1300	129026	AGGCTCACACCTGTATAGCTCAGCTATATTCAGCACTTTGGAGGCGAGGAGG	129085	QY	1301	AGATGACTTGGCTTGGCTTGGTTCAGAGCTTGGGCAACATGTTGAGACCTTGCTC	1360	129086	CAGATCACCTGAGTTCAGGCTGGAGACCGGCTTGGCAACATGTTGAAATCTGCTC	139145	QY	1361	CACAAACAAACAAACAAACAAACAAATCATTTACCTGGTACCATGGGTACCAAG--GTAC	1418	129146	TACTAAAAATACAAA--AATTAGCTGGGCATGTTGGGAGGCTGTAAATCCAGCTGTT	129204	QY	1419	ATAGAAATGACTCAGCAGATATGTTGCTCTCTCTACTTGTGGAGGCGGCTTATAC	1478	129205	GGGAGCTGAGGAGGAGATCGTTTGAACCCAGGAGGCGGCTTACAGTGAAGCGAGA	129264	QY	1479	TGAGTAAGCAATAGAGGAGGAGATATATATCTTAAATGAGAGGTACAGATTTGAGAG	1538	129265	TTGGCCACCGACTCCAGCTGGGCGACAAAGACAGCTGTCCCTCCCAAAAAA	129324	QY	1539	CAAAACAGGCGACAGGATATGACAGGTTAAAGAGGAAATCAGGAGGCTTCTCAG	1598	129325	AAAGTTCACTTGACATTTGAAGTGAACCCAGTGAATCCAGGATGCTGTCA--CCTCAAG	129382	QY	1599	AGAAGGTGACATTTAAGCCGACATGAAGATGAACAGTGTAGTTTACCAAGGATGGGA	1658	129383	GACTGAGCTTATCATCTGCAAGATTTCTTCCAAATTAAGGTCACTCGCATTCACA	129442	QY	1659	TGGAAGGGGTGAGAGTGAAGGAGGAGGAACTGCAGGATCATAGGCTTAGACAGG	1718	129443	GGTAGACATCTCATTTGAGGAGTGAAGGACAAACTTCACTCACTGCTCAGCCTCT	129502	QY	1719	GATCTGAGCCCTTGAAGAGTGAAGACAGCAGGAGTTCGTTAGTGGGTTAAGTAAC	1778	129503	TCCACCATTTTGTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	129562	QY	1779	AAAGCTGAGAGCCAGGAAATCCCTGCTCATGAGGCGCTGTGAGTCACTCAGAGTGT	1838	129563	CTGGAGCCCATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	129619	QY	1839	TGGGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1898	129620	TAAATCAGAGCATCTCCCTCGCCATGATGATGATGATGATGATGATGATGATGATGATGAT	129679	QY	1899	TTGGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1958	129680	CCAGGACTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	129737	QY	1959	TGATAGGAGGATCCAGGCTGATGGGAGGCTGGGAGTTCGCGGTGATGACACAGGCT	2018	129738	ATCTTCCAGCAGCATCTTGTGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	129795

RESULT 15	BD176843	DNA	linear	PAT 18-MAR-2003
BD176843	BD176843	79528 bp		
LOCUS				
				Db
				64926
				TCTTAA
				CAGTATT
				TAAGTCT
				CCAGGCC
				AGGCGTG
				GCTCATG
				CTGTG
				TATATCC
				CAGCAC
				64985

QY 2935 TTGGGAGGCTGAGGTGGTGGATCACAAGGTGAGGATCGAGACCATCCTGGCTAACA 2994
 |||||
 Db 64986 TTGGGAGGTGAAGGTGGACGGATCACCAGGTGAGGATCGAGACCATCCTGGCTAACA 65045
 |||||
 QY 2995 CAGTGAA 3001
 |||||
 Db 65046 CGGTGAA 65052
 |||||

Search completed: September 17, 2004, 14:58:13
 Job time : 11445.9 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 06:15:30 ; Search time 1107.52 Seconds
(without alignments)
11511.184 Million cell updates/sec

Title: US-10-669-693-3_COPY_1980_4980

Perfect score: 3001

Sequence: 1 cttcttcctccctaccctccccc.....atcctcggtacacagtgaa 3001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3001	100.0	17752	6	AAD45439 Human dru
2	257	8.6	79528	6	AAL50814 Human can
C 3	253	8.4	16747	4	AAL90152 Human dig
C 4	253	8.4	16747	4	AAL80326 Human imm
C 5	253	8.4	16747	4	AAL86604 Human imm
C 6	253	8.4	16747	4	AAL36989 Human mus
C 7	253	8.4	16747	7	ABX59977 CDNA enco
C 8	247.6	8.3	110096	6	ABN95044 Gene #154
9	247.2	8.2	2981	5	ABAL17151 Human ner
10	247.2	8.2	5041	5	ABAL17153 Human ner
C 11	245.8	8.2	110000	6	ABQ74964 5 Continuation (6 of
C 12	245.6	8.2	96593	9	ADC85367 Mouse BLM
C 13	245.6	8.2	96594	8	ADA02888 Human BLM
C 14	245.6	8.2	96594	9	ADB72626 Human BLM
C 15	245.6	8.2	99500	6	ADB72626 Human BLM
C 16	244.8	8.2	10500	6	ABK13086 Human REC
C 17	244.8	8.2	23130	4	AAD41740 Human imm
C 18	242	8.1	5952	9	AAL77045 Human imm
19	241.8	8.1	154465	6	ADB54002 PCNA geno
20	241.8	8.1	158245	6	AAD28763 Human AKA
21	241.8	8.1	161425	4	AAD28762 Human AKA
22	241.8	8.1	162025	4	AAL02340 Human AKA
23	241.8	8.1	162025	4	AAL02339 Human AKA
				6	AAD28758 Human AKA

24	241.8	8.1	162025	6	AAD28759	Aad28759 Human AKA
25	240.8	8.0	17327	2	AAQ44278	Aaq44278 Serglycin
26	240.6	8.0	185371	6	ABT10718	Abt10718 Human bre
C 27	240.2	8.0	98829	8	AAD55116	Aad55116 Human ACE
C 28	240	8.0	154902	6	ABQ88198	Abq88198 Human ost
29	238.8	8.0	8624	6	AAD44437	Aad44437 Human coa
C 30	238.8	8.0	10953	5	AAS29204	Aas29204 Genomic s
C 31	238.8	8.0	10953	6	ABS68344	Abs68344 Human DNA
C 32	238.8	8.0	10953	9	ADC25466	Adc25466 Human CDN
C 33	238.8	8.0	22645	4	AAK89727	Aak89727 Human dig
C 34	238.8	8.0	22645	4	AAK89727	Aak89727 Human dig
C 35	238.8	8.0	22645	4	AAK89727	Aak89727 Human rep
C 36	238.8	8.0	22645	4	AAK89727	Aak89727 Human rep
C 37	238.8	8.0	22645	4	ABL97878	Abi97878 Human tes
38	238.8	8.0	113604	8	ADA42607	Ada42607 Human tes
39	238.8	8.0	113604	8	ADA42607	Ada42607 Human IQT
40	238	7.9	357	4	AAK76716	Aak76716 Human imm
C 41	238	7.9	14874	4	AAK89726	Aak89726 Human dig
C 42	238	7.9	14874	4	AAK89726	Aak89726 Human dig
C 43	238	7.9	14874	4	ABL97877	Abi97877 Human tes
44	237.8	7.9	12452	4	AAL37624	Aal37624 Human mus
45	237.8	7.9	12452	5	ABA21062	Aba21062 Human ner

ALIGNMENTS

RESULT 1					
AAD45439					
ID	AAD45439 standard; DNA; 17752 BP.				
XX					
AC	AAD45439;				
XX					
DT	27-DEC-2002 (first entry)				
XX	Human drug-metabolising enzyme encoding gene.				
DE					
XX					
KW	Human; drug-metabolising enzyme; cytochrome p450; therapy; cytostatic;				
KW	Parkinson's disease; myotonic dystrophy; developmental defect; enzyme;				
KW	cancer; chromosome 19; gene; ds.				
XX					
OS	Homo sapiens.				
XX					
PH	Key				
FT	CDS	Location/Qualifiers			
FT		1999..13652			
FT		/*tag= a			
FT		/product= "Human drug-metabolising enzyme"			
FT	exon	1999..2167			
FT		/*tag= b			
FT	intron	2168..5236			
FT		/*tag= c			
FT	variation	replace(2226, GCT)			
FT		/*tag= d			
FT	variation	/*note= "Single nucleotide polymorphism"			
FT		replace(3081, TGC)			
FT		/*tag= e			
FT	variation	/*note= "Single nucleotide polymorphism"			
FT		replace(3788, T)			
FT		/*tag= f			
FT	variation	/*note= "Single nucleotide polymorphism"			
FT		replace(3979, TC)			
FT		/*tag= g			
FT	variation	/*note= "Single nucleotide polymorphism"			
FT		replace(5056, TC)			
FT		/*tag= h			
FT	variation	/*note= "Single nucleotide polymorphism"			
FT		replace(5213, GAC)			
FT		/*tag= i			
FT	variation	/*note= "Single nucleotide polymorphism"			
FT		5237..5386			
FT	exon	/*tag= j			
FT		5387..5919			
FT	intron	/*tag= l			

FT variation replace(5508, CTG)
FT /*tag= k
FT /note= "Single nucleotide polymorphism"
FT variation replace(5857, GAT)
FT /*tag= m
FT /note= "Single nucleotide polymorphism"
FT exon 5920..6080
FT /*tag= n
FT intron 6081..6166
FT /*tag= o
FT exon 6167..6636
FT /*tag= p
FT variation replace(6385, CT)
FT /*tag= q
FT /note= "Single nucleotide polymorphism"
FT intron 6637..13087
FT /*tag= r
FT variation replace(6813, CT)
FT /*tag= s
FT /note= "Single nucleotide polymorphism"
FT variation replace(7853, T)
FT /*tag= t
FT /note= "Single nucleotide polymorphism"
FT variation replace(12973, GCA)
FT /*tag= u
FT /note= "Single nucleotide polymorphism"
FT variation replace(13012, GAC)
FT /*tag= v
FT /note= "Single nucleotide polymorphism"
FT variation replace(13072, TCG)
FT /*tag= w
FT /note= "Single nucleotide polymorphism"
FT exon 13088..13261
FT /*tag= x
FT intron 13262..13441
FT /*tag= y
FT variation replace(13370, G)
FT /*tag= z
FT /note= "Single nucleotide polymorphism"
FT exon 13442..13648
FT /*tag= aa
FT variation replace(13682, CT)
FT /*tag= ab
FT /note= "Single nucleotide polymorphism"
FT variation replace(14631, CT)
FT /*tag= ac
FT /note= "Single nucleotide polymorphism"
FT XX US2002076774-A1.
PN
XX
XX
XX 20-JUN-2002.
XX
XX 27-DEC-2000; 2000US-00748127.
XX
XX 21-JUN-2000; 2000US-0212840P.
XX
XX (YANC/) YAN C.
XX (SHAO/) SHAO W.
XX (MERK/) MERKULOV G V.
XX (DFRA/) DI FRANCESCO V.
XX (BEAS/) BEASLEY E M.
XX
XX Yan C, Shao W, Merkulov GV, Di Francesco V, Beasley EM;
XX WPI; 2002-673347/72.
XX P-PSDB; AAB45439.
XX
XX A new human drug-metabolizing enzyme is related to the cytochrome p450
PT superfamily and is useful to provide modulators to treat drug-
PT metabolizing enzyme-related disorders including cancer and Parkinson's
PT disease.
XX
XX Claim 4; Page 23-31; 49pp; English.

XX The invention relates to human drug-metabolising enzyme peptides related
CC to the cytochrome p450 superfamily. Binding agents of the peptide are
CC used to treat a disease or condition mediated by a human drug-
CC metabolising enzyme. Such diseases include Parkinson's disease, cancer,
CC myotonic dystrophy and developmental defects. The present sequence is
CC human drug-metabolising enzyme encoding gene. This gene is located at
CC chromosome 19
XX
SQ Sequence 17752 BP; 2729 A; 3091 C; 2940 G; 2836 T; 0 U; 6156 Other;
Query Match 100.0%; Score 3001; DB 6; Length 17752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTCTTCCTCCTACCCCCAGCTGAGTAAGTAAGTACGACCGGTGTTCCACATCTACCT 60
DB CTTTCTTCCTCCTACCCCCAGCTGAGTAAGTAAGTACGACCGGTGTTCCACATCTACCT 2039
QY 61 GGGACCTGGCGGCTGTGCTGCTGCTGGGAGGAGGCTGTGCGGAGGCCCTGGG 120
DB CTTTCTTCCTCCTACCCCCAGCTGAGTAAGTAAGTACGACCGGTGTTCCACATCTACCT 2099
QY 121 AGGTGAGGCTGAGGAGTTTCAGCGCGCGGGGAAACCGTAGCGATGCTGGAAGGACCTTTGA 180
DB AGGTGAGGCTGAGGAGTTTCAGCGCGCGGGGAAACCGTAGCGATGCTGGAAGGAGCTTTGA 2159
QY 181 TGGCCATGTAAGTCAAGGCTGTAGGCCCTCGCTGAGCGTCCAGCTGCGACCACTTACTGG 240
DB TGGCCATGTAAGTCAAGGCTGTAGGCCCTCGCTGAGCGTCCAGCTGCGACCACTTACTGG 2219
QY 241 TGTGTGACCTTTGCACATGGCTTAGTCCCTCTGTTGCTCTCATCTGTCATTAATGAGTGATA 300
DB TGTGTGACCTTTGCACATGGCTTAGTCCCTCTGTTGCTCTCATCTGTCATTAATGAGTGATA 2279
QY 301 ACAGTGCCCATCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCAAACACTTTGGGAGCG 360
DB ACAGTGCCCATCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCAAACACTTTGGGAGCG 2339
QY 361 GAGGTGGGTGGATCACTTCAGGTTCAGAGTTTGAGACCAAGCTGGCCCAATGTTGAAC 420
DB GAGGTGGGTGGATCACTTCAGGTTCAGAGTTTGAGACCAAGCTGGCCCAATGTTGAAC 2399
QY 421 CCTGTCTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 480
DB CCTGTCTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2459
QY 481 GATACCTGGAGGTTGAGGCAGAGAAATCGCTTGAACCCCGGAGGAGATGTTGCAATGA 540
DB GATACCTGGAGGTTGAGGCAGAGAAATCGCTTGAACCCCGGAGGAGATGTTGCAATGA 2519
QY 541 ACCAAGACTGTGCCACTGCCATCCAGTCTGGGCAACAGAGTGGCTCCATCTCAACAA 600
DB ACCAAGACTGTGCCACTGCCATCCAGTCTGGGCAACAGAGTGGCTCCATCTCAACAA 2579
QY 601 ACAACAAAAGCAGTGGCCATCATGTAGGATTGAGTTGAGTGGAGGACTGAGCCTTG 660
DB ACAACAAAAGCAGTGGCCATCATGTAGGATTGAGTTGAGTGGAGGACTGAGCCTTG 2639
QY 661 TGCAAAAGTGAGCACTCACTAATCACCAGGTTGTAGTATCATGATCAATCAATGATCC 720
DB TGCAAAAGTGAGCACTCACTAATCACCAGGTTGTAGTATCATGATCAATCAATGATCC 2699
QY 721 AGGTAAGCCCTCAGGGTTTCAAGAGATCCGAGCGCTTTCAAGTGTCTGGGATTGTT 780
DB AGGTAAGCCCTCAGGGTTTCAAGAGATCCGAGCGCTTTCAAGTGTCTGGGATTGTT 2759
QY 781 GGGCAAGCCCTCGAATAAAGAAACAGTCTCTGTATTACACAGAAAGCAGAGGCCCA 840
DB GGGCAAGCCCTCGAATAAAGAAACAGTCTCTGTATTACACAGAAAGCAGAGGCCCA 2819
QY 841 TGCTGGGTGCTGCCAGAACTCAGTAGTAATTAAGACAGCACCGGTGCTGCTCCCGAGC 900
DB TGCTGGGTGCTGCCAGAACTCAGTAGTAATTAAGACAGCACCGGTGCTGCTCCCGAGC

PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 23-AUG-2000; 2000US-0227182P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 06-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 08-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 12-SEP-2000; 2000US-0232081P.
 PR 14-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 21-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0251989P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-502630/55.

Polynucleotides encoding digestive system antigens, useful for diagnosing, treating, preventing and/or prognosing disorders of the digestive system, particularly cancer and cancer metastases.

Disclosure; SEQ ID NO 3728; 986pp; English.

The present invention provides the protein and coding sequences of a number of human digestive system antigens. These can be used in the diagnosis, treatment and prevention of digestive system disorders, including cancer, Meckel's diverticulum, bacterial or parasitic infections, appendicitis, Hirschsprung's disease, chronic colitis or ulcerative colitis. The present sequence is a genomic DNA fragment encoding a digestive system antigen of the invention

Sequence 16747 BP; 4203 A; 3717 C; 3840 G; 3787 T; 0 U; 1200 Other;

Query Match 8.4%; Score 253; DB 4; Length 16747;

Best Local Similarity 27.2%; Pred. No. 1.5e-55;

Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;

QY 301 ACAGTCCCATCATCAGCGGGTGAGTGGCTAGTGCCTGAATCCCAACACTTTGGAGGCG 360

Db 16152 AAAATGAGTTAGGCTGGCGGGTGCTCAGGCTATATATCCCAACACTTTGGAGGCC 16093

QY 361 GAGGTGGTGGATCATCTTGGAGTTCAGAGCTTGAGCCAGCCTGGCCAAACATGGTGAAC 420

PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 41416; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I/
XX amino acid sequences given in AAM92170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 16747 BP; 4203 A; 3717 C; 3840 G; 3787 T; 0 U; 1200 Other;
SQ
Query Match 8.4%; Score 253; DB 4; Length 16747;
Best Local Similarity 27.2%; Pred. No. 1.5e-55;
Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;
QY 301 ACAGTGCCCATCAGCGGTCAGTGCTGCTGAAATCCCAACATTTGGGAGGCG 360
16152 AAAATGAGTTAGGCTGGCGCGGTGGCTCAGSCCTATAATCCCAACATTTGGGAGGCC 16093
Db
QY 361 GAGGTGGTGGATCACTTGAGGTGAGGAGTTCGAGACCGCTGGCCCAACATGGTGAAC 420
16092 GAGGCTGGTGGATCACTTGAGGTGAGGAGTTCGAGACCGCTGGCCCAACATGGTGAAC 16033
QY 421 CCTGTCTCTACTAAATAA-TAAATAATTAGCTGGGATGCTGTGGTACCTGTATATCC 479
16032 TCTGTCTCTACTAAATAATTCAAAAATTAGCTGGGATGCTGTGGTACCTGTATATCC 15973
QY 480 AGATACTTGGGAGGTTGAGCGAGGAGATCGCTTGAACCCGGAGGAGCAGATGTTGCAGTG 539
15972 AGCTACTTGGGAGGCTGAGGCGAGAGATTGCCTGAACCCAGAGATGAGGTTGTAGTG 15913
QY 540 AACCAAGACTGTGCCACTGCACTCCAGTCTGGGCAACA-GAGTGAGCCTCCATCTCAAAC 598
15912 AGCCAAGACTGCGCCACTGCACTCTAGCTGGGCAACAGAGCGGAACTCCGCTCAAAA 15853

RESULT 6
AAL36989/c
ID AAL36989 standard; DNA; 16747 BP.
XX
AC AAL36989;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3354.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatocellular; antidiabetic; antiinflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO20015367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001338.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226581P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229387P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.

17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including musculoskeletal
PT cancers and also for testing and detection e.g. diagnosis.
XX
PS Example 2; SEQ ID NO 3354; 781pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (AB03087-AB04109) associated with the musculoskeletal system useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. The genes are isolated from a range of human tissues
CC disclosed in the specification. The nucleic acids, proteins, antibodies
CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 16747 BP; 4203 A; 3717 C; 3840 G; 3787 T; 0 U; 1200 Other;
Query Match 8.4%; Score 253; DB 4; Length 16747;
Best Local Similarity 27.2%; Pred. No. 1.5e-55;
Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;
QY 301 ACAGTCCCATCAGCGGGTGCAGTGTAGTGCCTGAATCCCAACATTTGGAGGCG 360
Db 16152 AAAATGAGTTTAGGCTGGGCGGTGGCTCAGGCCTATATCCCAACATTTGGAGGCC 16093
QY 361 GAGGTGGGTGCATCAGTGTAGGTCAGAGTTCAGACACCGCTGGCCAAATGGTGAAC 420
Db 16092 GAGGCTGGTGGATCAGTTGAGGTTCAGAGTTAGAGACCGCTGGCCAAATGGTGAAC 16033
QY 421 CTTGTCTCTACTAAAAATA-TAAAAATTAGCTGGGCATGTGTGGTCTCTGTAATCCC 479
Db 16032 TCTGTCTCTACTAAATTACAAAAATTAGCTGGGCATGTGTGGGCTGCTGTAATCCC 15973
QY 480 AGATCTGGGAGGTTGAGCAGAGAGATCGTTGAACCGCGGAGGAGAGATTTGCAGTG 539
Db 15972 AGCTACTGGGAGGCTGAGGAGAGAAATTCCTGAACCAAGAGATGGAGTTGTAGTG 15913
QY 540 AACCAAGATGTGCCACTGCAGTCTCAGTCTGGGCAACA-GAGTGAGCCTCCATCTCAAAC 598
Db 15912 AGCCAAAGATGCGCCACTGCACTCTAGCTGGGCAACACAGCGGGAACCTCCGCTCAAAA 15853

QY 599 AAACAAACAAAAGACAGTCCCATCATGTAGGATTGAGTGTAGTGGAGGACTGAGCCT 658
Db 15852 AAAAAAATAAATTTGGAGGGAGATGAGATTGTAATGTAAATATATAGACATCATTTTA 15793
QY 659 TGTGCAAAAGTGAGCACTCACTAATCACCAGGTTGTAGTATCAGTGTATACCATCATGAT 718
Db 15792 TAGGAATAGTGAGTCTCTGAAGACCTCTATGATGGTTAATTTGGGGAAGATGTTCAAT 15733
QY 719 CCAGTAAAGCCCTGAGGGTTCAGAAAGATGCCGAGCGCTTTCAAGTGTCTGGGATG 778
Db 15732 TCTGAGTCAGAAAAATATTTACAGACAATTTTGTCTGTCAGTTCAGAGGATGGAGAACAGA 15673
QY 779 GTGGCAAGCCCTCGAATAATAGAAAAGTCTCTGTATTATACAAAGAGAGAGAGGCC 838
Db 15672 ATGGAGGAGAGAAAGGCAATTCAGAAAGTGTGGCATTTATTTCAAGGTGAGAGTAAGACAG 15613
QY 839 CATGCTGGTGTCTGCCAGGAATCAGTAGTAATAAGACAGCACCGGTGCTCTCCCA 898
Db 15612 TATCAGTAGAGTGAAGAAGGTAAAGACACATGCCAAAGATTAATCAAGGATAAATA 15553
QY 899 GCGCACCTAGCCAGTGGGAAACAGACTCCACACAGTCCCGAGCCAGAGTGTCTCAGG 958
Db 15552 TTGAAGGTAATTTGTAAGTTATAGTTTGGGTGGAATTTCTGGTTTACCAATTCGTG-G 15374
QY 959 GCCAAGATGGGAGACACGGGGAG-----AAAGTCAGGTGGGATGGGAGAGGG 1008
Db 15492 GCGAGTTTGGAGATAAGATCCTGTTGAAATTTCTGTGTTAAGGCTTTCTGAGGATTC 15433
QY 1009 TCAGGCAAGAGGGGTTCAGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAG 1068
Db 15432 TCAGTGGAATTTTGTAGGTTATAGTTTGGGTGGAATTTCTGGTTTACCAATTCGTG-G 15374
QY 1069 GAGTACTGACCCCGCATGTTTGGTGAGGAGATTCAGGAAGTCTTCTCTGGAAGAGAGG 1128
Db 15373 TAGCTGAAGCAAGAACAAAGAAATATCTTGATCAAGGGTATGTAAGTTCATGAGAAAGG 15314
QY 1129 CTGTGCGAGTGCAGCTCATAGATGAGTGGGAGGGTGTTC--AGGCAAGAAAGACCA 1185
Db 15313 AAAAAGAGCTTAAATGAGAAAGGGTGAATAAACAGGAGACAGAAATGGAAGAGCAT 15254
QY 1186 GCACCTACAAAAGCATGACTTTGAGAGAGCATTCATCCATTCATGATCAATTTTCAG 1245
Db 15253 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 15194
QY 1246 ACTGGGCACTGCTGCTCATGCTGTAATCCAGCATTTTGAAGAGGCTGAATGGGAGGAT 1305
Db 15193 NNN 15134
QY 1306 GACTTGAGCCTAGGCATTTGTGCAAGCCTGGGCAACATGGTGAGACCCCTGCCTCCCAA 1365
Db 15133 NNN 15074
QY 1366 AACAAACAAACAAACAAACAAATCATTTATCTGCTACCATGGGTACAGGTACATAGAAA 1425
Db 15073 NNN 15014
QY 1426 TGACTCAGGAGATATGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1485
Db 15013 NNN 14954
QY 1486 AGACAATAGGGAGGGAATAATCTTAAATGAGAGGTACAGATTGAGAGCAACAC 1545
Db 14953 NNN 14894
QY 1546 AGGGCAGCGCATATCTACGAGGTAAAGGGAATCAGGGAAGGCTTCTCAGAGAGGT 1605
Db 14893 NNN 14834
QY 1606 GACATTTAAGCCGGGACATGAAGGATGAACGAGTTAGTTCAACAGGATGGGATGAAAG 1665
Db 14833 NNN 14774
QY 1666 GGGTGAGTGTGGAGGAGGGAACCTGCAGGATCATAGGCCTAGACGGGGATCTCTG 1725

[illegible]

QY 1486 AGACAATAGAGGAGGGAATATAATCCTAAATGAGAGGTACAGATTGTGAGCAAAACAC 1545
Db 14953 NNN 14994
QY 1546 AGGCACAGGCATATGTACGAGGTAAAGAGGAATCAGGGAAGGCTTCTCAGAGAAGT 1605
Db 14893 NNN 14834
QY 1606 GACATTTAAGCCGGGACATGAAGGATGAACGAGTTAGTTTCAACAAAGGATGGATGGAAG 1665
Db 14833 NNN 14774
QY 1666 GGGTCAGAGTGATGAGGACAGAGGAACTGCAGGATCATAGGCCCTAGACAGGGGATCCTG 1725
Db 14773 NNN 14714
QY 1726 ACGCCCTTGAGGAAGTGAGAGAAGACCAGCGCAGTCGTAGTGGGTTAAAGTAAACAAAGCTG 1785
Db 14713 NNN 14654
QY 1786 AGAAGCCAGGGAATCCCTGGTCATGACGGGCTGTGAGTCAGTCAGTCAGTGTTTGGGCT 1845
Db 14653 NNN 14594
QY 1846 TTTGTTTCTCTGGGACAGTCGATTTTAAGCAGGGAACAGCTGATTTCAGAGTTGGGAA 1905
Db 14593 NNN 14534
QY 1906 GATCCTGTGTTGCTCTGAAGGGGATGAATCTGGAGCTAGGAGCCAGGCTGATAGG 1965
Db 14533 NNN 14474
QY 1966 GAGGATCCAGGGTGATGGGAGGCTGGGAGGTCGCGGTGATGAGCACGGGCTGGGCCA 2025
Db 14473 NNN 14414
QY 2026 GGGGATGGGAGAGGAGTAAATTGGAGAGGCTGGGCTCTGGCCGAGGAATGATGG 2085
Db 14413 NNN 14354
QY 2086 TGGGCTGAACAGGAGAGGAGAGATGCTTAGGCCACTTTGGAACAACAGTAGGGCAAGA 2145
Db 14353 NNN 14294
QY 2146 CAGGAGACCCAAAGGGAGAGTCCCAAGAGACCAGCAGGCTGGCATTTGACAGGGAA 2205
Db 14293 NNN 14234
QY 2206 GGTCTGTCTGGACAGGTGCTTTGGATAAGGGAGGAAATGGTGCACTTCCATCCTCCTC 2265
Db 14233 NNN 14174
QY 2266 CCTCTCTGTCTCAACTTAACATACATGGGACAGGACCCAGTGGGACTCCATAATGA 2325
Db 14173 NNN 14114
QY 2326 TGGGATGGGTGGATGAGAGAAGAGGAGGAACAACACTTTCATTCCTGTTATTT 2385
Db 14113 NNN 14054
QY 2386 ACAGAACAGCCAGGTCGCGTGTCTAGCTTGGCATTTAGCACCTTTGGGAGGCTGAGGT 2445
Db 14053 NNNNNNNNNNNNNNNNGTGGCTCAGGCCTATAATCCCAACACTTTGGGAGGCTGAGGT 13994
QY 2446 GGGTGAATTAACCTCAGTCAGGAGTTCAAGACAGCCCTAGACACAGTAGAGAACC--- 2502
Db 13993 GGGTGAATTTCTTGAGCCAGGGGTTCAAGACCAGCTTGGGCAATATGTTGAACCCCTGC 13934
QY 2503 ---CATCTCTCAAGATATAAATAGCTGGCGTAGTGGCATATGCTGCTGTAATCCCA 2559
Db 13933 CTGACTCTATAAAAAAATACAAAATTAGCCAGAGGTGGTGTGATCTGTAGACCCA 13874
QY 2560 GCTAGTCGGGAAGCTGAGGCAGGAGAATCGTTTGAACCCGAGAGGAGGCTTGGGTGA 2619

Db 13873 GCTACTTAGGAGGCTGAGGAGGAGGATTGCTTGAGTCCGAGGTCAAGGCTGCAGTGA 13814
QY 2620 GCTGAGATCGGTCATTGCACTCCAGCTCGGTGACAAAGCAAGACCTCGTCTCAATAAT 2679
Db 13813 ACAATGATCAAGCCACTGCACTCCAGCTTGAGCGACGTGAGCACTGCTCTCAATAAAA 13754
QY 2680 AATAATAATTACAAAACAGAA 2700
Db 13753 CCCAAAACCAACAAACAA 13733
RESULT 8
ABN95044/c
ID ABN95044 standard; DNA; 110096 BP.
XX ABN95044;
AC
XX 13-AUG-2002 (first entry)
DT
XX Gene #1542 used to diagnose liver cancer.
DE
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX Homo sapiens.
OS
XX WO200229103-A2.
PN
XX 11-APR-2002.
PD
XX 02-OCT-2001; 2001WO-US030589.
PF
XX 02-OCT-2000; 2000US-0237054P.
PR
XX (GENE-) GENE LOGIC INC.
PA
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
PI WPI; 2002-426119/45.
XX
XX Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX
PS Claim 1; SEQ ID NO 1542; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 110096 BP; 25112 A; 29361 C; 28188 G; 27435 T; 0 U; 0 Other;

Query Match 8.3%; Score 247.6; DB 6; Length 110096;
Best Local Similarity 47.3%; Pred. No. 8.7e-54;
Matches 1149; Conservative 0; Mismatches 1224; Indels 57; Gaps 11;
QY 287 CAATGAGTGATAACAGTCGCCATCAGCCGGGTGCGAGTGGCTAGTGCCTGAATCCCAA 346
Db 58480 CAACAAAAGAAAACAGTAAGGCTGCTGGGACGGTGCACGCTGTAATCCAG 58421

QY 347 CACTTTGGAGCGGAGGTGGGTGATCACTTGGCTCAGGATTCGAGACGACGCTGGC 406
Db |||||
58420 CAATTTGGAGGCCAGGACGGTGGATCACCTAGGTCAGGAGTTCAAGACGAGCCTGGC 58361
QY 407 CAACATGTTGAACCCCTGTCTCTACTTAAATAATAAAATTTAGCTGGGCATGGTGGTGGC 466
Db |||||
58360 CAACATGTTGAACCCCTATCTCTACTTAAATAATAAAATTTAGCTGGGCATGATGGCAGG 58301
QY 467 TACCTGTAATCCGATACCTTGGGAGGTTGAGGACGAGGAATCGCTTGAACCCGAGGAGGC 526
Db |||||
58300 TGCCTGTAAATCCGAGTACTCAGGAGGCTGAGATGGGAAATCACTTGAACCTGGAGAT 58241
QY 527 AGATGTTTGCAGTGAACCAAGACTGTGCCACTGCCTCCAGTCTGGGCAACAGAGTGAGCC 586
Db |||||
58240 AGTGGTTGCACTGAGCCGAGATCTGGCAGCTGCATCTCATCTGGCAGCTGAGCGAGAC 58181
QY 587 TCCATCTCAAAACAAACAAACAAAGAGCTGCCCATCATGTAGGATTCAGTGAATTTGAGTG 646
Db |||||
58180 CTCGTCTAAAAAATAAAAAAATAAAAAAAGGCAAAATTTATCCAGGCATGGTTGGTAG 58121
QY 647 AGGACTGAGCCTTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 706
Db |||||
58120 GCACCTGTAAATCCGAGTACTCAGGAGGCTGAGTGGGAGAAATTCCTTGAACCCAGGAGG 58061
QY 707 ACCATCAATGATCCAGGTAAAGCCCTGAGGGTTCAGAAAGATGCCGAGCGCTTTCAGG 766
Db |||||
58060 TGGAGGCTGCACTGAGCCAGATCATGCCACTGCCTCCAGCTGGGCGACAGAGTGAGA 58001
QY 767 TGTGGGGATTTGGGCAAGCCCTCGAATAATAGAAACAGTCTCTGTATTACACAGA 826
Db |||||
58000 CCCCTTCTCAAAAAAGAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 57941
QY 827 AAGCAGAGGCCATCTGTGGGTGCTGCCAGAACTCAGTAGTAATTAAGCAGCAGCCGGT 886
Db |||||
57940 GGG-TGGAGGCCAGGGCCCTGGAAAGGCCCTGTATGCTGTGTGAAGGAATTTAGACTTGTAT 57882
QY 887 GCTGCTTCCCAGCGCACCTAGGCCAGTGGGGAACAG-----ACTCACCACAGT 938
Db |||||
57881 CTGTGCTGGTGGGATGGTTTGTAGCTTAACAACAGGAGTAGGAATCAGATTACAGT 57822
QY 939 CCCAGCCAGAGTGTGAGGCGCAAGATGGGAGACACGGGGA---GAAAGGTCAAGGTG 995
Db |||||
57821 GTAGAATGTCACTAGGGCTGACACTGTGGCTCAGCTTGTATCTCAGCACTTTGGGG 57762
QY 996 GATGGGGAGGGTCAAGGCAAGAGGGTCAAGGCCAGCTGAGGAAAGCCCTGGGACTG 1055
Db |||||
57761 CCAGGCAGGATTTTGTAGGCCAGGAGTTTGAACACAGCTTGGGTAAACATAGCAAGACC 57702
QY 1056 TAGGAATTTAGAGGAGGTACCTGACCCCGCATGTTTGTGAGGAGATTCAGGAAGTCT 1115
Db |||||
57701 TCCATCTATACAGAAAAATTAACAAAAATTTAGCCAGGAGTAGTGGTACCCGCTGTAGT 57642
QY 1116 CTGGAAGAGAGGCTGTGAGGCTGAGA--CTCATAGATGAGTGGGGAGGGTGTTCAG 1173
Db |||||
57641 CCTAGATATTCAAGAGCACAGACGAGAGATCACTTGAACCTGGGGGAGTTCAAAGTTG 57582
QY 1174 GCAGAAGACCAACCTACAAAAGCATGACTTTGAGAAGCATTCATCCATTCACATG 1233
Db |||||
57581 CAGTAATTTGATGTGATCTCAGCTTGGCAATGTAGCAAAACCCCTGCCCTTCAAAAGAAA 57522
QY 1234 ATGAATTTTTCAGACTGGGACGCTGGCTCATGCTGTATCTCCAGCACTTTTGAAGGCTG 1293
Db |||||
57521 AAAAAAGTTCAGCCGGTGGCTGAGTCAAGCTGTATCTCCAGCACTTTTGGAGGCGG 57462
QY 1294 AATGGGAGGATGACTTGAAGCTTGAAGCTTTGTGACAAGCCTGGGCAACATGTTGAGACC 1353
Db |||||
57461 AGACGGCGGAT--CATGAGGTGAGCAGATTGAGACCATCTCTGGCTTAACATGTTGAACC 57404
QY 1354 CTGCTCTCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1413
Db |||||
57403 CCGCCTCTACTAAAAATACAAAAAATTTAGCCAGGTGTGGTGGCGGCGCTGTGTCCCA 57344

QY 1414 GGTACATAGAATAAGTACTCAGCAGATATGGTGT---CCTCTCCTACTGTGGAGAGGGCG 1470
Db |||||
57343 GCTACTCGGAGGCTGAGGCGAGGAATGGATGAACCCGGGAGCGAGAGCTTTCAGTGA 57284
QY 1471 GCTTATCTCAGTAAGACAATAGAGGAGGGAATATATATCTTAAATAGAGGTTACAGA 1530
Db |||||
57283 GCCGAGATTGTGCCACTTGCATCTCAGCCTGGGGACAGAGGAGACTCCATCTCAAAAA 57224
QY 1531 TTTGAGAGCAAAACACAGGGCACAGGCATATGTACAGGGTAAAGAGGGAATCAGGGAAG 1590
Db |||||
57223 AATATATATATATAAAGGTGAGCGTGGTGGCTCATGCCCTATATATCCAGCACTTTGG 57164
QY 1591 CTTTCTCAGAGAGTGTGATTTAAGCCGGGACATGAAGGATGAACGAGTTAGTTACCAA 1650
Db |||||
57163 AGGCCGAAGCAGGCGGA---TCACCTTGAGGTTCAGAGTTTGAGACCAGCTTGGCCAA 57107
QY 1651 GGATGGGATGGAAGGGGTGAGGTGATGGAGGCGAGGGAATTCAGGATCATAGGCT 1710
Db |||||
57106 GGTGAGACCTGTCTCTACTTAAATAATAAATTTAGCCGGCGTGGTGTGTGTGCTG 57047
QY 1711 AGACAGGGGATTCCTGACGCCCTTCAGGAAGTGAAGAGAACCCAGCGAGTCTGTAGTGGT 1770
Db |||||
57046 TAGTCCCAAGCTACTCAGAGGCTGGGCGACAGAAATGACTCGAGCCCAAGGAGGTGAGGT 56987
QY 1771 TAAGTAACAAAGCTGAGAACGCCAGGGAATCCCTGTGTATGCAAGGCTGTGAGTCACT 1830
Db |||||
56986 GGCAGTGAGTGAAGTTCACCACTCCAGCCTTGAGCGACACTCCAGCCTGTCACTCCAGC 56927
QY 1831 CAGAGTCTTTGGGCTTTTCTTCTGGGAGAGTTCGATTTTAAAGCAGGGAACAGCTGT 1890
Db |||||
56926 CTGAGCACCGAGTGAGAGTCTGTCTCAAAAAAATAAATAATAAATAAATAAATAAATAA 56876
QY 1891 ATTGAGAGTTGGGAAGTCCCTGTGGTGTCTGCTGAAAGGGATGAAACTTGAGGCTAGGA 1950
Db |||||
56875 TTTTTCACAAAAGAAAAATTTCTTACAAAAGAAAAAATTTTTCACATAAATTTATT 56816
QY 1951 GCCCAGGCTATAGGAGGATCCAGGCTGATGGGAGGCTGGGAGGTCCGGGCTGATGGA 2010
Db |||||
56815 ACAAATTTTTCACAAAATAAATAAAGTAATAAATAAATAAATAAATAAATAAATAAATAA 56756
QY 2011 CCAGGCTGGGCGCAGGGGATGGGAGGAGGAGTAATTCGGAGAGGCTGGGCTCTGG 2070
Db |||||
56755 GAAGTGAAGTTATAGAGTCTTAATGATGTGTAGATTTAGAGACATTAAGAGGTAGGGC 56696
QY 2071 CCAGGAATGATGGTGGCTGAAACAGGAGAGGAGAGATGCTTAGGCCACTTTTGAAC 2130
Db |||||
56695 CGGCGATGTGGCTCAGCCTGTAAATCCAGCACTTTGGGAGGCTGAGGCGAGGTAGA-- 56639
QY 2131 ACAGTGGGCAAGGACAGGAGACACCCAGGGGAAGTGCCCAAGAGACACAGCAGGCTG 2190
Db |||||
56638 TCACGAGGTGAGGATGGAGACCCAGCCTGGCCAGATGTTGAACCCCTCTCTACTAA 56579
QY 2191 GCATTCGACAGGGAAGTCTGTCTGGAGCAGGTGTCTTGATTAAGGAGGAAAAATGTGC 2250
Db |||||
56578 AATACAAAAATTTAGCTGGTGGTGGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 56519
QY 2251 AGTTCCATCT 2310
Db |||||
56518 CGAGGCAAGAGATCGC-----TTGAAACCGGAGTTGGAGGTTGAGGTTCAGTGCCTGAGAT 56465
QY 2311 GGAATCCATAATGATGGATGGGTGGTGGATGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2370
Db |||||
56464 CGCTCCACTGCACTCTAGCTTGGG-----CAACAAGAGCAAGACTCTGT 56422
QY 2371 TCATCTGTGTTATTTACAGAACAGGCGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2430
Db |||||
56421 CTCAAAAAATAAATAAATAAAGGCTGGGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 56362
QY 2431 TTGGGAGGCTGAGGTGGTGGATTTACCTCAGGTTCAGAGTTTCAAGACCAAGCTTAGACAAC 2490
Db |||||
56361 TTGAGAGGCGGAGGAGGCGGATCACCTGAGGTTCGGAGTTTGAGACCAAGCTTGAACCAAC 56302
QY 2491 GTAGAGAAACCCCATCTCTACTGAAGATATAAATAATGAGTGGGCGTAGTGGCATATGCT 2550

Db	56301	ATGGAGAAATCCGCTCTCTACTATAAAATACAAAATTCGCCAGTGTGGTACATGCCT	56242	PR	22-AUG-2000;	2000US-0226681P.
				PR	22-AUG-2000;	2000US-0226868P.
				PR	22-AUG-2000;	2000US-0227182P.
Qy	2551	GTAATCCAGCTAGTCGGGAAGCTGAGCAGGAGAAATCGTTGAACCGAGAGGAGG	2610	PR	30-AUG-2000;	2000US-0227009P.
				PR	01-SEP-2000;	2000US-0228924P.
Db	56241	GTAATCCAGCTATTTCGGGAGCTGAGGCAAGAGAAATCGCTTGAACCGGAGGAGG	56182	PR	01-SEP-2000;	2000US-0229343P.
				PR	01-SEP-2000;	2000US-0229344P.
Qy	2611	TTGGCGTGAGCTAGATCGTGCCATTGCACTCCAGCCTGGGTGACAAAGCAAGCTCGT	2670	PR	01-SEP-2000;	2000US-0229345P.
				PR	05-SEP-2000;	2000US-0229509P.
Db	56181	TTGCAGTGAGCGGAGATCGCGCATTTGCATCCAGCCTGGGCAATAGAGTGAACCTCCA	56122	PR	05-SEP-2000;	2000US-0229513P.
				PR	06-SEP-2000;	2000US-0230437P.
Qy	2671	CTCAATAATAATAATAATTACAAAACAGAA	2700	PR	06-SEP-2000;	2000US-0230438P.
				PR	08-SEP-2000;	2000US-0231242P.
Db	56121	TCTGAAAAAAAATAATTAATAATAAAAA	56092	PR	08-SEP-2000;	2000US-0231243P.
				PR	08-SEP-2000;	2000US-0231244P.
				PR	08-SEP-2000;	2000US-0231413P.
				PR	08-SEP-2000;	2000US-0231414P.
				PR	08-SEP-2000;	2000US-0232080P.
				PR	08-SEP-2000;	2000US-0232081P.
				PR	12-SEP-2000;	2000US-0231968P.
				PR	14-SEP-2000;	2000US-0232397P.
				PR	14-SEP-2000;	2000US-0232398P.
				PR	14-SEP-2000;	2000US-0232399P.
				PR	14-SEP-2000;	2000US-0232400P.
				PR	14-SEP-2000;	2000US-0232401P.
				PR	14-SEP-2000;	2000US-0233063P.
				PR	14-SEP-2000;	2000US-0233064P.
				PR	14-SEP-2000;	2000US-0233065P.
				PR	21-SEP-2000;	2000US-0234223P.
				PR	21-SEP-2000;	2000US-0234224P.
				PR	25-SEP-2000;	2000US-0234997P.
				PR	25-SEP-2000;	2000US-0234998P.
				PR	26-SEP-2000;	2000US-0235484P.
				PR	27-SEP-2000;	2000US-0235834P.
				PR	27-SEP-2000;	2000US-0235836P.
				PR	29-SEP-2000;	2000US-0236327P.
				PR	29-SEP-2000;	2000US-0236367P.
				PR	29-SEP-2000;	2000US-0236368P.
				PR	29-SEP-2000;	2000US-0236369P.
				PR	29-SEP-2000;	2000US-0236370P.
				PR	02-OCT-2000;	2000US-0236802P.
				PR	02-OCT-2000;	2000US-0237037P.
				PR	02-OCT-2000;	2000US-0237038P.
				PR	02-OCT-2000;	2000US-0237039P.
				PR	02-OCT-2000;	2000US-0237040P.
				PR	13-OCT-2000;	2000US-0239935P.
				PR	13-OCT-2000;	2000US-0239937P.
				PR	20-OCT-2000;	2000US-0240960P.
				PR	20-OCT-2000;	2000US-0241785P.
				PR	20-OCT-2000;	2000US-0241786P.
				PR	20-OCT-2000;	2000US-0241787P.
				PR	20-OCT-2000;	2000US-0241808P.
				PR	20-OCT-2000;	2000US-0241809P.
				PR	20-OCT-2000;	2000US-0241826P.
				PR	01-NOV-2000;	2000US-0244617P.
				PR	08-NOV-2000;	2000US-0246474P.
				PR	08-NOV-2000;	2000US-0246475P.
				PR	08-NOV-2000;	2000US-0246476P.
				PR	08-NOV-2000;	2000US-0246477P.
				PR	08-NOV-2000;	2000US-0246478P.
				PR	08-NOV-2000;	2000US-0246523P.
				PR	08-NOV-2000;	2000US-0246524P.
				PR	08-NOV-2000;	2000US-0246525P.
				PR	08-NOV-2000;	2000US-0246526P.
				PR	08-NOV-2000;	2000US-0246527P.
				PR	08-NOV-2000;	2000US-0246528P.
				PR	08-NOV-2000;	2000US-0246532P.
				PR	08-NOV-2000;	2000US-0246609P.
				PR	08-NOV-2000;	2000US-0246610P.
				PR	08-NOV-2000;	2000US-0246611P.
				PR	08-NOV-2000;	2000US-0246613P.
RESULT 9						
ABAI17151						
ID	ABAI17151	standard; DNA; 2981 BP.				
XX	AC	ABAI17151;				
XX	XX					
DT	23-JAN-2002	(first entry)				
XX		Human nervous system related polynucleotide SEQ ID NO 9482.				
DE						
XX		Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;				
KW		immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;				
KW		antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;				
KW		antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;				
KW		antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;				
KW		antiparasitic; cardiant; immune disorder; cardiovascular disorder;				
KW		neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.				
XX						
OS		Homo sapiens.				
XX						
PN	WO200159063-A2.					
XX						
PD	16-AUG-2001.					
XX						
PF	17-JAN-2001;	2001WO-US001334.				
XX						
PR	31-JAN-2000;	2000US-0179065P.				
PR	04-FEB-2000;	2000US-0180628P.				
PR	24-FEB-2000;	2000US-0184664P.				
PR	02-MAR-2000;	2000US-0186350P.				
PR	16-MAR-2000;	2000US-0189874P.				
PR	17-MAR-2000;	2000US-0190076P.				
PR	18-APR-2000;	2000US-0198123P.				
PR	19-MAY-2000;	2000US-0205515P.				
PR	07-JUN-2000;	2000US-0209467P.				
PR	28-JUN-2000;	2000US-0214886P.				
PR	30-JUN-2000;	2000US-0215135P.				
PR	07-JUL-2000;	2000US-0216647P.				
PR	07-JUL-2000;	2000US-0216880P.				
PR	11-JUL-2000;	2000US-0217487P.				
PR	11-JUL-2000;	2000US-0217496P.				
PR	14-JUL-2000;	2000US-0218290P.				
PR	26-JUL-2000;	2000US-0220963P.				
PR	26-JUL-2000;	2000US-0220964P.				
PR	14-AUG-2000;	2000US-0224518P.				
PR	14-AUG-2000;	2000US-0224519P.				
PR	14-AUG-2000;	2000US-0225213P.				
PR	14-AUG-2000;	2000US-0225214P.				
PR	14-AUG-2000;	2000US-0225266P.				
PR	14-AUG-2000;	2000US-0225267P.				
PR	14-AUG-2000;	2000US-0225268P.				
PR	14-AUG-2000;	2000US-0225270P.				
PR	14-AUG-2000;	2000US-0225447P.				
PR	14-AUG-2000;	2000US-0225477P.				
PR	14-AUG-2000;	2000US-0225757P.				
PR	14-AUG-2000;	2000US-0225758P.				
PR	14-AUG-2000;	2000US-0225759P.				
PR	18-AUG-2000;	2000US-0226279P.				

CC associated (CA) genes from the 50 tables given in the specification. The
 CC CA proteins are secreted, transmembrane or intracellular proteins. The
 CC recombinant nucleic acids are useful for screening for drug candidates
 CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
 XX ADC85514 represent CA genes of the invention.

SQ Sequence 96593 BP; 27724 A; 19526 C; 19631 G; 29712 T; 0 U; 0 Other;

Query Match 8.2%; Score 245.6; DB 9; Length 96593;
 Best Local Similarity 48.2%; Pred. No. 2.8e-53;
 Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;

QY 314 GCCGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGAGGGGAGGTGGTGGAT 373
 DB 17837 GCCGGCGTGGTGGCTTATGCTCTTAATCCCGACCTTTGGAGGCGGAGACGGGCGAT 17778

QY 374 CACTTGAGTCAAGGAGTTCGACACCATGCTGCCCAACATGGTGAACCCCTGCTCTACTA 433
 DB 17777 CACTTGAGTCAAGGAGTTCGACACCATGCTGCCCAACATGGTGAACCCCTGCTCTACTA 17718

QY 434 AAAATATAAAATTAGTGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 493
 DB 17717 AAAATATAAAATTAGTGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 17658

QY 494 TTGAGCAGGAGAAATCGCTTGAAACCGGAGGAGGAGTGTTCAGTGAACCAAGACTGTGC 553
 DB 17657 TTGAGCAGGAGAAATCGCTTGAAACCGGAGGAGGAGTGTTCAGTGAACCAAGACTGTGC 17598

QY 554 CACTGCACCTCCAGTCTGGGCAACAGAGTGAAGTCCATCTCAAAACAAACAAACAAAGC 613
 DB 17597 CACTGCACCTCCAGTCTGGGCAACAGAGTGAAGTCCATCTCAAAACAAACAAACAAAGC 17541

QY 614 AGTGCCCATCATGTAGATTTAGATTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 673
 DB 17540 AGTGCCCATCATGTAGATTTAGATTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 17487

QY 674 CTCACATAATC-ACCAGGTTGATGATCAGTGAATCAATCAATGATCCAGGTGAAGCCCT 732
 DB 17486 CTCACATAATC-ACCAGGTTGATGATCAGTGAATCAATCAATGATCCAGGTGAAGCCCT 17427

QY 733 GAGGGTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGGATGGTGGGCAAGCCCTC 792
 DB 17426 GAGGGTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGGATGGTGGGCAAGCCCTC 17367

QY 793 GAATAATAGAAACAGTTCTCTGTATTAACAGAAAGAGAGGAGGAGGAGGAGGAGGAGG 852
 DB 17366 GAGGGTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGGATGGTGGGCAAGCCCTC 17307

QY 853 CCAGGAACCTCAGTAGTAACTTAGACAGACCGGTGCTCTCCACAGCGCACCTAGGCCA 912
 DB 17306 GAGGGTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGGATGGTGGGCAAGCCCTC 17247

QY 913 GTGGGAAACAGACTCACCACACAGTCCAGGCCAGAGTGGTCAAGGCCCAAGATGGGAA 972
 DB 17246 GAGGGTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGGATGGTGGGCAAGCCCTC 17187

QY 973 GCACGGGAGAAAGGTCAAGGTGGATGGGAGGAGTCAAGGAGGAGGAGGAGGAGGAGGAG 1032
 DB 17186 GAGGGTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGGATGGTGGGCAAGCCCTC 17127

QY 1033 GGTGAGGGAAGCCCTGGGAGTCTAGGAATTTAGAGGAGGTACTCTACCGCGCATGTTG 1092
 DB 17126 GAGGGTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGGATGGTGGGCAAGCCCTC 17073

QY 1093 GTGAGGAGATTCAGGAAGTCTTCTGGAAGAGAGGCTGTCGGAGCTGAGACTCATAGA 1152
 DB 17072 GTGAGGAGATTCAGGAAGTCTTCTGGAAGAGAGGCTGTCGGAGCTGAGACTCATAGA 17016

QY 1153 TGAGTGGGAGGAGTGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1212
 DB 17015 TGAGTGGGAGGAGTGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16973

QY 1213 AAGCAATTCATCCATTCAGTGAATTTTTCAGACTGGGAGGAGGAGGAGGAGGAGGAGGAG 1272

DB 16972 AAATATTGAGCTTTAAAGAAATAATT--GACTAGATGAGTGGCTCATGCTGTAA 16915
 QY 1273 TCCAGCACTTTGGAAAGCTGAATGGGAGGAGTGAATCTTGGAGCTAGGCAATTTGTGACAG 1332
 DB 16914 TCCAGCACTTTGGAGGCAAGGAGGAGTATCGCTTTGAGCCAGGAGTTTGTGAGACCA 16855

QY 1333 CCTGGGCAACATGGTGAAGCC--TGCCTCCCAAAACAAACAAACAAACAAACAAATCAT 1390
 DB 16854 CCTAGGCAACAAAGCAAGACCCAGTATCTTACAAATAATAAAAGTTAGCCAGCATGTT 16795

QY 1391 TATACCTCGTACATGGGTACCAGGTACATAGAAATGACTCAGGAGGAGTATGTTGCTCT 1450
 DB 16794 GGTACACTCTCTAGTCTTCACTCTCGGAGGCTGAGGCAAGGATCATCTGAGCCCA 16735

QY 1451 TCTTACTGTGGAGAGGCGGCTTATCTGAGTGAAGCAATAGAGGAGGAGGAGGAGGAGGAG 1510
 DB 16734 GGAGTTAGAGGCTGAGTGAAGCAAGTATCTTCTGACCTCCAGCTTGGATAGCTGAG 16675

QY 1511 CCTAAATGAGAGGTACAGATTTGAGAGCAACACAGGSCACAGGCATATGTACGAGGCT 1570
 DB 16674 CAAGACCTCTGAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 16615

QY 1571 AAAGAGGGAATCAGGAGAGGCTTCTCAGAGAAGTGACATTTAAGCC-----GGGACATG 1625
 DB 16614 AAAGAGGGAATCAGGAGAGGCTTCTCAGAGAAGTGACATTTAAGCC-----GGGACATG 16555

QY 1626 AAGGATGAACAGTGTAGTTTCAACAGGATGGATGGAAGGGGTGAGATGATGAGGAGGCA 1685
 DB 16554 AAGTAAAGCCAGTCAACAGGACAAATATCATGATTTTCACTGCAACAGGAGTATCTA 16495

QY 1686 GAGGGAATGCAAGGATCATAGGCTTAGACAGGAGGATCTTGCAGCCCTTCAGGAAGTGAGA 1745
 DB 16494 GGCAGGATCATAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16435

QY 1746 GAAGACCCAGCGAGTGTAGTGGGTTAAGTAAACAAAGCTGAGAAGCCAGGGAATCCCTG 1805
 DB 16434 TGGGAGGTTGTTAATCAAAAGATATAAAGTTTTCAGTAATGCAAAATGAACGATTTCTAGA 16375

QY 1806 GTATGAGGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1865
 DB 16374 GATCTGCCATACAACATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 16316

QY 1866 TCGATTTTAAAGCAGGAGCAAGCTGATTCAGAGTTGGGAGGATCTTGTGCTGCTGCTGCTG 1925
 DB 16315 TTAAGAAACAAATTTTAAAGGATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 16256

QY 1926 AAGGGGATGAAATCTGAGGCTAGGAGCCAGGAGTGAAGGAGGATCCAGGAGTGAAGGG 1985
 DB 16255 GCTAGAGCGAGTGGGACAAATCATAGTTCACTGCACTCAGACTCCTAAGCTCAAGTGA 16196

QY 1986 AGGCTGGAGGCTCGCGGTGATGGACAGG--GCTGGGGCCAGGGGATGGGAGGAGGAGG 2042
 DB 16195 TCTCTGCTGCTGAGCTCTGGAGATCTGGAATCAGAGTGCATGATCATCATCTCAACT 16136

QY 2043 AGTAATTTGGAGAGGCTCGGGCTCTGGCCGAGGAATG-----GATGTGGGCTGAAAC 2096
 DB 16135 AATTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 16076

QY 2097 AGGAGGAGGAGATGCTTTAGGCCACTTTTGGAAACACAGTAGGGCAAGGAGGAGGAGGAG 2156
 DB 16075 CTGGCTCAAGAGATCTTGTCTTCTCAGGCTCCCAAGTGTGGGATTCAGAAAGTGAAG 16016

QY 2157 CAAGGGAGGAGTCCCAAGAGACACAGCAGGCTGGCATTTGGAGGAGGAGGAGGAGGAGGAG 2216
 DB 16015 TACCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 15956

QY 2217 AGCAGGTGTCTTGGATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2267
 DB 15955 ACAACGTGGAGGAGTCTCAAAACATATAACGCAAACTAGAAAGAGTCAAGTAGAATGCA 15896

QY 2268 TCTCTCTCAACCTCTTAAACTATCATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2327

QY 1153 TGAATGGGGAGGGTGTTCAGGCGCAAGAACAGCAGACCTCAAAAAGCATGACTTTGAGAG 1212
 Db 17015 AGAGGAAATGGGAGTTTCTTTTACAGTGGGTATACGCA-----TAGAATG 16973
 QY 1213 AAGCATTCATCCATCACTCACTGATGAATTTTCAGACTGGGCGCTGGCTCATGCTGTAA 1272
 Db 16972 AAATATTAGCCTTTTAAAGAAATAAATT--GACTAGATGAAGTGGCTCATGCTGTAA 16915
 QY 1273 TCCAGCAGCTTTGGAAGGCTGAATGGGAGGATGACTTTGAGCCTAGGCCATTTTGTGACAAG 1332
 Db 16914 TCCAGCAGCTTTGGGAGGCCAAGGAGGAGTATCGCTTGGCCAGGAGTTTGGAGACCAG 16855
 QY 1333 CTTGGGCAACATGTGTGAGACC--TGCCTCCACAAAACAAAACAAACAAACAAATCAT 1390
 Db 16854 CCTAGGCAACAAAGCAAGCACCAGTATCTACAAAATAAATAAAGTTAGCCAAAGCATGT 16795
 QY 1391 TATACCTGTATACATGGGTACCAGGTACATAGAAATGACTCAGGCGAGATATGGTGTCCCTC 1450
 Db 16794 GGTACACTCTGTAGTCTCACCCTTCTCGGAGGCTGAGCGCAGAGGATCATCTGAGCCCA 16735
 QY 1451 TCCTACTGTGGAGAGCGGGCTTATATCTGCGATGAGA CAATAGAGGGAGGGAATATAAT 1510
 Db 16734 GGAGTTAGAGGCTGCAGTGAGCAACGATCATTTCTACTGCACTCCAGCCTGGATAGCTGAG 16675
 QY 1511 CCTAAAATGAGAGGTACAGATTTGAGAGCAAAACACAGGCGCACAGGCATATGTACGAGGCT 1570
 Db 16674 CAAGACCTCTGTCTGAAAAATAATAAATAACAATACAAATTAAGTTAAATAATAAATAA 16615
 QY 1571 AAAGAGGGAATCAGGGAAGGCTTCTCAGAAAGGTGACATTTAAAGCC-----GGGACATG 1625
 Db 16614 AAAAAAATCTGCAACATGAGATGACACAGATGAACCTTGAAGCCATTAAAGCAAGTG 16555
 QY 1626 AAGGATGAACGATTTAGTTTACCAAGATGGGATGGGATGGAAGGGGTGAGATGATGGAGGCA 1685
 Db 16554 AAGTAAGCCAGTACACAGGCAAAATACATACATGATTTCCACTGCAACAGGATATCTAAATA 16495
 QY 1686 GAGGGAACCTGCAGGATATAGGCTAGACAGGGGATCCTGAGCCCTTGGAGGAGTGAGA 1745
 Db 16494 GCAAACTCATGAAGCAAGAGGGGCACAGTGTGGCAGGGTAGGGCTTTCAGGAAA 16435
 QY 1746 GAAGACCGCGCAGTCTGATGGGTTAAGTAAACAAAGCTGAGAAAGCCAGGGAATCCCTG 1805
 Db 16434 TGGGAGTTGTTAATCAAAAAGTATAAGTTTCAGTAATGCAAAATGAACGATTTCTAGA 16375
 QY 1806 GTCATGCGAGGCTGTGATGATCAGTCAGAGTGTGGGCTTTGTTTCTCTGGGAGCAG 1865
 Db 16374 GATCTGCCATACAAATTTGTGGC-TGTGTGGTACTATACATGTAAAACTCTGTTAAGAG 16316
 QY 1866 TCGATTTTAAAGCAGGCAACAGCTGTATTCAGAGTTGGGAGATCCTGTGTTGCTGCCTG 1925
 Db 16315 TTAAGAAAACAATTTAAAAAATTTTTTTTTTAGACAAAGTCTTGCTCTGCTCACTCAG 16256
 QY 1926 AAGGGATGAACTGGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGGTGTGGG 1985
 Db 16255 GCTAGAGCGCAGTGGCACAATCATAGTTCACTCAGCCTCAGACTCCTTAAGCTCAAGTGA 16196
 QY 1986 AGGCTGGAGGCTCGGCTGATGGACAGG--GCTGGGGCCAGGGGATGGGAGGAAG 2042
 Db 16195 TCCTCTGCTCAGCCTCTGAGAAATCTGGAATCTGAGGTGTCATGATCATGCTCAACT 16136
 QY 2043 AGTAATTTGGAGAGGCTGGGGCTCTGGCCGAGGAATG-----GATGTGGGCTGAAAC 2096
 Db 16135 AATTTTTAAATTTTATAGACAAAGTCTCCATTTGTGCCAGACTGCTTGAATC 16076
 QY 2097 AGGAGAGGAGAGATGCTTAGGCCACTTTTGGAAACAGATAGGGCAAGAGCAGAGACACC 2156
 Db 16075 CTGGCTCAAGAGATCTTGCTTCTCAGCCTCCCAAGTCTGGGATTAACAAGTGGAGC 16016
 QY 2157 CAAGGGAGTGCCCAAGAGACCAACAGAGCTGGCATTTGGACAGGGAAGGTCTGCTGG 2216
 Db 16015 TACCATGCCAGCCAAAATTTACTATTAAAAAATAAAGAAAAAAGAAAAAATAATATCT 15956

QY 2217 AGCAGGTGTCTTGGATAAGGGAGGAAAATGGTGC-----AGTTCATCTCTCTCC 2267
 Db 15955 ACAACGTGAAGAATCTCACAAACATATAACGCAAACTAGAAAGACTCAGTAGAATGCA 15896
 QY 2268 TCTCTGTTCAACCTCTAAACTACATGGGGCACAGGACCCAGTGGGACTCCATAAATGATG 2327
 Db 15895 TATAGTATTATCCCATTTACAGAGCTCAAAACCCAGGCAAACTGAGCTATATTATTAT- 15837
 QY 2328 GGATGGGTGATGGAAGGAAGGAAGGAAGAAACAACCTCTTCAATCATCTGCTTATTAC 2387
 Db 15836 AGATGTATAGATCAGGGCAAACTATGAAGACAAGAAAAAGCAATTATCACAAAAATGGGAC 15777
 QY 2388 AGAACAGGCCAGGTGCGGTGCTCAGCTTGCCTTGCCTTCTAGCACTTTGGGAGGCTGAGGTG 2447
 Db 15776 AATAGGCCAGAGCGGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGCCGAGCGAG 15717
 QY 2448 GTGGAATTACCTCAGGTTCAGGAGTTCAAGACCAAGCCTTAGACAACGTAGAGAAAACCCATCT 2507
 Db 15716 GCAGATTACCTGAGGTTCAGGAGTTTAAGACCAAGCCTGGCCAAACATGATGAACCCCGTCT 15657
 QY 2508 CTACTGAGAT--ATAAATTAGCTGGGGTAGTGGCATATGCTGTAAATCCAGCTAGT 2565
 Db 15656 CTACTAAAAATAGAAAAAATTAGCAGCCATGCTGGCAGCCCTGTAAATCCAGCTACT 15597
 QY 2566 CGGGAAGCTGAGCGAGGAGAAATCGCTTGAACCCGAGAGCAGAGGTTGCGGTGAGCTGAG 2625
 Db 15596 TGGGAGGCTGAGCGAGGAGAAATGCTTGAACCCAGGAAGCAGAGGTTGCAAGTGGCCGAG 15537
 QY 2626 ATCGTGCCATTTGACTCCAGCCTGGTGACAA-AGCAAGACCTCTGCTCAATAATAATAA 2684
 Db 15536 ATCGGCGCACTGCACTCCAGAAATGGGCAACAAGAGCAAAATCCGCTCTCAAAATAATAA 15477
 RESULT 14
 ADB72626/c
 ID ADB72626 standard; DNA; 96594 BP.
 XX
 AC ADB72626;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human BLM gene.
 XX
 KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
 XX
 OS Homo sapiens.
 XX
 PN WO2003008583-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 26-DEC-2001; 2001WO-US051291.
 XX
 PR 02-MAR-2001; 2001US-00798586.
 PR 23-OCT-2001; 2001US-00004113.
 PR 08-NOV-2001; 2001US-00052482.
 PR 30-NOV-2001; 2001US-0097722.
 PR 20-DEC-2001; 2001US-00034650.
 XX
 PA (SAGR-) SAGRES DISCOVERY.
 XX
 PI Morris DW, Engelhard EK;
 XX
 DR WPI; 2003-239337/23.
 XX
 PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.
 XX
 PS Claim 1; SEQ ID NO 454; 2304pp; English.
 XX
 CC The invention relates to a novel recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the 660 sequences fully defined

CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a human gene of the invention.

```
XX SQ Sequence 96594 BP; 27725 A; 19526 C; 19631 G; 29712 T; 0 U; 0 Other;

Query Match      8.2%; Score 245.6; DB 9; Length 96594;
Best Local Similarity 48.2%; Pred. No. 2.8e-53;
Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;

QY 314 GCCGGTGCAGTGGCTAGTCCCTGAAATCCCAACACTTTGGAGCGGAGGTGGGTGGAT 373
DB 17837 GCCGGGCTGGTGGCTTATGCCGTGAATCCCNAGCACTTTGGAGGCCGAGACGGCAGAT 17778
QY 374 CACTTGAGGTGAGGAGTTGAGACCAAGCTGGCCCAACATGFTGAAACCCCTGCTCTACTA 433
DB 17777 CACCTGAGGTGAGGAGTTTGGACCACTCTGGACCAACATGGCAAAACGCTCTCTACTA 17718
QY 434 AAAATATAAAATTAAGCTGGGCATGGTGGTACCTGTATCCAGATACCTGGGAGG 493
DB 17717 AAAATACAAAATTAAGCAGGCATGATGGTGACACCTTAAACCCAGCTATTCAGTAGG 17658
QY 494 TTGAGGCGAGGAATCGCTTGAACCCGGGAGCGAGATGTTGAGTGAACCAAGACTGTGC 553
DB 17657 CTGAGACATGAGATCTCTTGAACCCAGCTGGTAGAGGTTCAGTGGACCAAGATCATAC 17598
QY 554 CACTGCATCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAAAACAAACAAACAAAGC 613
DB 17597 CACTGCATCCAGCTGGGCGACAGTGAGACTCCTTCTCAAAACAAACAAACAAACAA 17541
QY 614 AGTGCCCATCATGTAGGATTCAGTATGATGAGTGAAGGCTGAGGCTGAGCCCTGTGCAAGTGACCA 673
DB 17540 -----CTTCTCTGGGTAGACAAAGGGCTGGGGGTGAGTACTTAATCACCATGGCCAA 17487
QY 674 CTCATAATC-ACCAGTGTGATGATCAGTGAACCATCAATGATCCAGGTAAAGCCCT 732
DB 17486 TTATTTAATCAATCATGTCTAGGTAAAGGAGCCCTACGTAAACCCAGAGGATAGGGCT 17427
QY 733 GAGGGTTCAGAAAGATCCGAGCGCTTTCAAGGTGCTGGGGATGTTGGTGGCAAGCCCTC 792
DB 17426 CAGGGAGCTTCTAGGTGGTGAGAACATGTAGAGATGGTGGGAGCGTGTGCTCGCTGGAG 17367
QY 793 GAATAATAGAAACAGTCTCTGTATTACAAAGAAAGACAGAGGCCCATGTGGTGGCTG 852
DB 17366 AGGCATGGAAGTCCACACCCCTTCCCAATACCTTGTCTGTGTATTTCTTCCAGCTG 17307
QY 853 CCAGGAACCTCAGTAGTAACCTAAGACAGCACCGGTGCTGCTTCCCGAGCGCACCTAGGCCA 912
DB 17306 GCTGCTTCTGAGTTGTTTATAATAAGCTGATAATCTAGTAAGGAACTGTTCTGTGACTG 17247
QY 913 GTGGGAAACAGACTACCAACAGTCCAGCCAGAGTGTGAGGCCAAGATGGGGAA 972
DB 17246 AATGTGGAGACTTACCACATTTGTCTGTAGTCTGAAAGTGGGGGTCTGGAATCGCGGTG 17187
QY 973 GCACGGGGAAGAGGTTCAGGTGGATGGGAGGGGTTCAGGCAAGAGGGTTCAGGGCCA 1032
DB 17186 GACAGGAGAGACTTAACATTTGGGGTCTGCATCAACCTGGCAGTGTAGTGTTCAGACTG 17127
QY 1033 GGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGACCCCGGCATGTTG 1092
DB 17126 AGCTAAATGTAGGACACCCAGTGTCTCTAAAGAACTGAACTATTGTGGGGAA----- 17073
QY 1093 GTGAGGAGATTCAGGAAGTCTTCTCGAAGAGAGGCTGTGGAGCTGAGACTCATAGA 1152
DB 17072 ----AAAAACACCAACATTTGGTGACAGAAGTGAGGAATTTATATATGGGTACAGAGTAG 17016
QY 1153 TGAGTGGGAGGTTTCCAGGCGAGAAGACCAGCACCTCAAAAGCATGACTTTGAGAG 1212
DB 17015 AGAGGAATCGAGTTTCTTTTACAGTGGTATACGCA-----TAGAATG 16973
QY 1213 AAGCATTCATCCATTCAACTGATGAATTTTTCAGACTGGGCAAGCGCTGGCTCATGCTGTAA 1272
```

```
DB 16972 AAATATTCAGCTTTAAAGAAATAAATTT--GACTAGATGAAGTGGCTCATGCTTGTAA 16915
QY 1273 TCCACGACTTTTGAAGGCTGAATGGGAGGATGACTTGAAGCTTAGGCAATTTGTGCAAG 1332
DB 16914 TCCACGACTTTTGGAGGCCAAGGCGAGGATTCGCTTGGAGCCAGGAGTTTGGAGCCAG 16855
QY 1333 CTGGGCAACATGTTGAGACCC--TGCCCTCCAAAAACAAACAAACAAACAAATCAT 1390
DB 16854 CCTAGGCAACAAAGCAAGACCCAGTATCTACAAAAATAAAAAAGTTAGCCAGCATGCT 16795
QY 1391 TATACCTGTCACCATGGGTGAGGATCCAGGTACATAGAAATGACTCAGGCAGATATGCTCTC 1450
DB 16794 GGTACACTCTGTGTCTCACCCTTCTCGGGAGGCTGAGGCAGAGGATCATCTGAGCCCA 16735
QY 1451 TCCTACTTGGAGAGGGGGCTTATCTGTCAGTAAGACAATAGAGGGAGGCAATATATAT 1510
DB 16734 GGAGTTAGAGGCTGCAGTGAGCAACGATCATCTACTGCACTCCAGCCCTGGATAGCTGAG 16675
QY 1511 CCTAAATGAGAGGTACAGATTTGAGACAAACACAGGGCAAGGCATATGTACAGGGT 1570
DB 16674 CAAGACCCCTGTCTGAAAAATAATAAAATACAATAACAATTAAGTTAAATAATAAAAAATTA 16615
QY 1571 AAAGAGGGAATCAGGAGGCTTCTCAGAGAGGTGACATTTAAGCC-----GGACATG 1625
DB 16614 AAAAAAATCCTGCAACATGAGATGACACAGATGAACCTTGAAGCCATTAAGCAAGTG 16555
QY 1626 AAGGATGAACGAGTTAGTTTACCAAGGATGGGATGGAAGGGGTGAGAGTGTATGAGGCA 1685
DB 16554 AAGTAAGCCAGTCCAGGACAAATACTATCATGATTTCCACTGCACGAGGTATCTAAATA 16495
QY 1586 GAGGAACTGCAGGATCATAGCCCTAGACGGGATCTGACGCCCTTTGAGGAAAGTGAGA 1745
DB 16494 GGCAAACTCATAGAAGCAAGAGGGGCACAGTGGTTGCCAGGTAGGGGCTTTCAGGAAA 16435
QY 1746 GAACACAGCCAGTCTGTAGTGGTAAAGTAACAAGCTGAGAGCCAGGCAAGATCCCTG 1805
DB 16434 TGGGGAGTTGTTAATCAAAAGTATAAAGTTTCAGTAATGCAAAATGAACGATTTCTAGA 16375
QY 1806 GTCATGAGGGCCCTGTGAGTCACGTACAGAGTGTGTTGGGCTTTTGTCTCTGGAGCAG 1865
DB 16374 GATCTGCCATACAACTTTGTGGC--TGTTGTGTACTATACATGTAATACTCTGTTAAGAG 16316
QY 1866 TCGATTTTAAAGCAGGAAACAGCTGTATTCAGAGTTGGGAAGATCCTGTGTTGCTGCTG 1925
DB 16315 TTAAGAAAAACAATTTAAAAAAATTTTTTTTAGAGACAAGTCTTGTCTGTCACTCAG 16256
QY 1926 AAGGGATGAACCTGGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGCTGATGGG 1985
DB 16255 GCTAGCGCAGTGGCAATCATAGTTCACTGCAGCCTCAGACTCTCTAAGCTCAAGTGA 16196
QY 1986 AGGCTGGAGGTCCGCGGTGATGGACCCAGG---GCTGGGGCCAGGGATGGGAGGAGG 2042
DB 16195 TCCTCTGCTCAGCCTCTGAGAACTCTGGAACCTACAGGTGCAATCATGCTCACT 16136
QY 2043 AGTAATTTGGAGAGCCCTGGGGCTCTGGCCGAGGAATG-----GATGGTGGGCTGAAC 2096
DB 16135 AATTTTTTAAATTTTTTATAGAGACAAGTCTCCCAATTTGTTGCCAGACTGCTCTTGAAC 16076
QY 2097 AGGAGAGGAGAGATGCTTAGGCCACTTTTGGACACAGTAGTAGGCAAGGACAGGACACC 2156
DB 16075 CTGGCTCAAGAGATCTTGTCTCAGCCTCCCAAAGTCTGGGATTACAGAGTGAGC 16016
QY 2157 CAAGGGGAAGTGCCCAAGAGACCAAGCAGGCTGGCAATTCGACAGGGAAGTCTGTCTGG 2216
DB 16015 TACCATGCCAGCCAAAATTTACTATTAAAAAAAATAAAAAAGAAAGAAACTATATCT 15956
QY 2217 AGCAGGTGCTTGGATAAGGGAGGAAATGGTGC-----AGTTCCATCTCTCTCCC 2267
DB 15955 ACAACGTGGAAGAACTCTCAAAACATAATAACGCAAACTAGAAAGACTCAGTAGAATGCA 15896
QY 2268 TCTCTGTTCAACCTCTAAACTACATCGGGGCAAGGACCCAGTGGGACTCCATAAATGATG 2327
```


Db 7977 CCTAGGCAACAAGCAAGACCCAGTATCTACAAAAATAAAGATTAGCCAGCATGGT 7918
QY 1391 TATACCTGGTACCATGGGTACCAGGTACATAGAAATGACTCAGGCAGATATGGTGTCTTC 1450
Db 7917 GGTAACACTCTGTACTCTCAGCTTCCTCGGGAGGCTGAGGCAGAAGGATCATCTGAGCCCA 7858
QY 1451 TCCTACTGTGGAGAGGGGGCTTATCTGCTAGTAAGACAATAGAGGGAGGGAATATAT 1510
Db 7857 GGAGTTAGAGGCTGCAGTGAGCAACGATCTCTACTGCACTCCAGCCTGGATAGCTGAG 7798
QY 1511 CCTAAATGAGAGGTACAGATTGTAGAGCAAAACACAGGCACACAGGCATATGTACAGGGT 1570
Db 7797 CAAGACCCCTGTCTGAAAAATAATAAATACATACAATTAAGTTAAATAATAAATAA 7738
QY 1571 AAAGAGGAATCAGGGAAGGCTTCTCAGAGAAGGTGACATTTAAGCC-----GGGACATG 1625
Db 7737 AAAAAAATCCTGCACATGAGATGACACAGATGAACCTTTGAAGCCATTAAAGCAAAAGT 7678
QY 1626 AAGGATGAACGAGTTAGTTTACCAGGATGGATGGAAGGGGTGAGAGTATGAGAGCA 1685
Db 7677 AAGTAAGCCAGTCAAGGACAAATACTACATGATTCCTACTGCAAGCAGGTATCTAAATA 7618
QY 1686 GAGGAACTGACGATCATAGCCCTAGACAGGGATCCTGACGCCCTTGAGGAAATGAGA 1745
Db 7617 GGCAACTCATAGAAGCAAGAGGGGCACAGTGGTTGCCAGGGTAGGGCTTTCAGGAA 7558
QY 1746 GAAGACCAGCGCAGTCTGTAGTGGGTAAAGTAACAAGCTGAGAAGCCAGGAAATCCCTG 1805
Db 7557 TGGGAGTTGTTATCAAAAAGTATAAGTTTACAGACAGAGTCTTGCTCTGTCACTCAG 7498
QY 1806 GTCATGAGGCGCTGTGAGTCACCTCAGAGTGTGTTGGGCTTTGTTTTCTCGGGAGCAG 1865
Db 7497 GATCTGCCATACAACATTTGGGC-TGTGTGTACTATACATGTAAGTCTGTTAAGAGG 7439
QY 1866 TCGATTTTACGAGGAAACAGCTGTATTCAGAGTTGGGAAAGATCCTGTGGTTGCTGCTG 1925
Db 7438 TTAAGAAAAACAATTTAAAAAAATTTTATAGACACAGGTCTTGCTCTGTCACTCAG 7379
QY 1926 AAGGGGATGAACCTGGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGGTATGGGG 1985
Db 7378 GCTAGAGCGAGTGGCACAATCATAGTTCCTACTGAGCCTCAGATCCCTAAGCTCAAGTA 7319
QY 1986 AGGCTGGAGGTCGGGCTGATGACACAGG---GCTGGGCCAGGGGATGGGAGGAGG 2042
Db 7318 TCCTCTGCTCAGCCTCTGAGAACTCTGGAACCTAGAGGTGATGACATCATGCTCACT 7259
QY 2043 AGTAATTGGGAGAGGCTGGGGCTCTGGCCGAGGAATG-----GATGGTGGGCTGAAAC 2096
Db 7258 AATTTTAAATTTTATAGACACAGTCTCCCAATGTTGCCAGAGTGGTCTTGAACCTC 7199
QY 2097 AGGAGAGGAGAGATGCTTAGGCCACTTTTGAACACAGTAGGSCAAGGACAGGACAGCACC 2156
Db 7198 CTGGCCTCAAGAGATCTTGCTTCTCAGCCTCCAAAGTCTGGGATTAAGAAAGTGAGC 7139
QY 2157 CAAGGGGAAGTCCCAAGAGACACACAGCAGGCTGGCAATGGACAGGAGAGGCTGTCTGG 2216
Db 7138 TACCATGCCAGCAAAATTTACTATTAAAAAAGAAAAAGAAAGAAAGAACTATATCT 7079
QY 2217 AGCAGGTGTCTGGATAAGGAGGAAAAATGGTGC-----AGTTCCATCCTCCTCCC 2267
Db 7078 ACAAGTGAAGAACTCTCAAAACATAAATACCAAACTAGAAAGACTCAGTAGAATGCA 7019
QY 2268 TCTCTCTTCAACCTCTAAACTACATGGGCGCAGGACCCAGGAGTGGGACTCCATAATGATG 2327
Db 7018 TATAGTATTATCCATTTACAGACTCAAAACAGGCAAAACTGAGCTATATTTAT- 6960
QY 2328 GGATGGTGGATGGAAGGAAGGAGGAAACAACTTTTCATTCCTGTTATTTAC 2387
Db 6959 AGATGTATAGATCAGGGCAAACTATGAGACAGAAAGCAATATCAAAAAATTTGGGAC 6900
QY 2388 AGAACAGGCGAGGTGGGTGCTCAGCTTGCCATTCTAGCATTTGGGAGGCTGAGGTGG 2447

Db 6899 AATAGGGCCAGAGCGGTGGCTCATGCTCTGTAAATCCAGCACTTTGGGAGGCCGAGCAG 6840
QY 2448 GTGGATTACCTCAGGTCAGGAGTTCAAGACCAAGCCTTAGACAACGTTAGAGAAACCCCATCT 2507
Db 6839 GCAGATTACCTCAGGTCAGGAGTTTAAAGACCAAGCCTGGCCAAACATGATGTAACCCGCT 6780
QY 2508 CTACTGAAGAT--ATAAAATTAGCTGGGCGTAGTGGCATATGCCCTGTATATCCAGCTACT 2565
Db 6779 CTACTAAAAATAGAAAAAATTAGCCAGCCTATGGTGGCAGGCGCTGTATATCCAGCTACT 6720
QY 2566 CGGGAAGCTGAGGCAGGAGAAATCGCTTGAACCCGAGAGGAGAGGTTGCGGTGAGCTGAG 2625
Db 6719 TGGGAGGCTGAGGCAGGAGAAATTGCTTGAACCCAGGAAGCAGAGGTTGCACTGAGCCGAG 6660
QY 2626 ATCGTGCAATTGCACTCCAGCCTGGGTGACAA-AGCAAGACCTCGCTCTCAATAATAATAA 2684
Db 6659 ATCGGCCACTGCACTCCAGAAATGGGCAACAAGACAAACTCCGCTCTCAAAATAATAA 6600

Search completed: September 17, 2004, 08:42:04
Job time : 1143.52 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 07:38:50 ; Search time 192.326 Seconds
(without alignments)
8659.273 Million cell updates/sec

Title: US-10-669-693-3_COPY_1980_4980
Perfect score: 3001
Sequence: 1 cttcttcctccctaccacc.....atcctggtaacacagtgaa 3001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	245.8	8.2	786431	US-09-751-389-3	Sequence 3, Appli
C 2	245.6	8.2	99500	US-09-798-096-10	Sequence 10, Appl
C 3	240.8	8.0	17327	US-07-906-871-15	Sequence 15, Appl
C 4	234	7.8	21968	US-09-851-985-3	Sequence 3, Appli
C 5	232.6	7.8	2839	US-09-061-702-1	Sequence 1, Appli
C 6	232.2	7.7	48763	US-09-916-204-3	Sequence 3, Appli
C 7	232	7.7	66804	US-09-740-041-3	Sequence 3, Appli
C 8	231.4	7.7	72928	US-09-009-913-1	Sequence 3, Appli
C 9	231.2	7.7	128779	US-09-497-855A-38	Sequence 38, Appl
C 10	230.6	7.7	174493	US-09-804-471A-3	Sequence 3, Appli
C 11	230.6	7.7	174493	US-10-238-709-3	Sequence 3, Appli
C 12	230.4	7.7	16063	US-09-301-052-3	Sequence 3, Appli
C 13	230.4	7.7	16063	US-10-020-121-3	Sequence 3, Appli
C 14	229	7.6	246240	US-08-724-394A-20	Sequence 20, Appl
C 15	229	7.6	246240	US-08-724-394A-21	Sequence 21, Appl
C 16	229	7.6	246240	US-08-724-394A-22	Sequence 22, Appl
C 17	228.8	7.6	90541	US-09-759-359A-3	Sequence 3, Appli
C 18	228.4	7.6	11288	US-08-646-301A-1	Sequence 1, Appli
C 19	228.4	7.6	11288	US-08-481-968A-4	Sequence 4, Appli
C 20	228.4	7.6	11288	US-08-154-712B-4	Sequence 4, Appli
C 21	228.4	7.6	15056	US-09-474-699-10	Sequence 10, Appl
C 22	228.2	7.6	685	US-09-183-266A-16	Sequence 16, Appl
C 23	228.2	7.6	53332	US-09-801-861-3	Sequence 3, Appli
C 24	228	7.6	116592	US-09-818-512-3	Sequence 3, Appli
C 25	227.6	7.6	45716	US-08-965-048-5	Sequence 5, Appli
C 26	227.6	7.6	45989	US-08-965-048-6	Sequence 6, Appli
C 27	227.6	7.6	128779	US-09-497-855A-38	Sequence 38, Appl

C 28	227.2	7.6	36651	4	US-09-738-894A-3	Sequence 3, Appli
C 29	227.2	7.6	36651	4	US-09-964-469-3	Sequence 3, Appli
C 30	227	7.6	36741	3	US-09-301-665-3	Sequence 3, Appli
C 31	226.6	7.6	70000	4	US-09-851-896-3	Sequence 3, Appli
C 32	226.2	7.5	399	4	US-09-621-976-13959	Sequence 13959, A
C 33	226.2	7.5	10380	3	US-09-077-354B-3	Sequence 3, Appli
C 34	226	7.5	497	4	US-09-621-976-3876	Sequence 3876, Ap
C 35	226	7.5	3627	4	US-09-323-873A-6	Sequence 6, Appli
C 36	226	7.5	99500	4	US-09-798-096-10	Sequence 10, Appli
C 37	225.8	7.5	36651	4	US-09-738-894A-3	Sequence 3, Appli
C 38	225.8	7.5	36651	4	US-09-964-469-3	Sequence 3, Appli
C 39	225.2	7.5	368	4	US-09-621-976-10146	Sequence 10146, A
C 40	225.2	7.5	2115	1	US-08-395-800A-7	Sequence 7, Appli
C 41	225.2	7.5	118067	4	US-09-497-855A-32	Sequence 32, Appl
C 42	224.6	7.5	62804	4	US-09-800-960-3	Sequence 3, Appli
C 43	224.6	7.5	62804	4	US-10-096-960-3	Sequence 3, Appli
C 44	224.4	7.5	283	4	US-08-579-445-26	Sequence 26, Appl
C 45	224.4	7.5	841	5	PCT-US93-06251-80	Sequence 80, Appl

ALIGNMENTS

RESULT 1
US-09-751-389-3/c
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C1001067
; CURRENT APPLICATION NUMBER: US/09/751.389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

Query Match	8.2%	Score	245.8	DB	4	Length	786431
Best Local Similarity	84.2%	Pred. No.	7.8e-57				
Matches	277	Conservative	0	Mismatches	52	Indels	0
Gaps	0						
QY	282	TCGTCAATGGAGTGATACAGTGCACATCAGCGGGTGCAGTGGCTAGTGGCTGCTGAAT	341				
Db	572672	TTTTTCACTACTGTTATTAACAGTACTAGAGCCAGGTGCTGTGGCTCAAGCTATAAT	572613				
QY	342	CCCAACACTTTGGAGCGGAGGTGGGTGATCACTTAGGTGAGGTGAGGATTCGAGACCAGC	401				
Db	572612	CCCAACACTTTGGAGCGGAGGTGGGTGATCACTTAGGTGAGGTGAGGATTCGAGACCAGC	572553				
QY	402	CTGGCCAAACATGGTGAACCCCTCTCTACTATAAATAATAAATAAATAAATAAATAAATAA	461				
Db	572552	CTGGCCAAACATGGTGAACCCCTCTCTACTATAAATAATAAATAAATAAATAAATAAATAA	572493				
QY	462	GTGCGTACCTGTAATCCAGATACCTTGGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG	521				
Db	572492	GGGTACCTGTGATCCAGCTACTTGGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	572433				
QY	522	GAGGAGAGTGTTCAGTGAACCAAGACTGTGCACCTGCACCTGCACCTGCACCTGCACCTGC	581				
Db	572432	GAGGAGAGTGTTCAGTGAACCAAGACTGTGCACCTGCACCTGCACCTGCACCTGCACCTGC	572373				
QY	582	GAGCCTCCATCTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA	610				

Db	572372	GAAACTCGTCTCAAAACAAACAAAA	572344	
RESULT 2				
US-09-798-096-10/c				
; Sequence 10, Application US/09798096				
; Patent No. 6399378				
; GENERAL INFORMATION:				
; APPLICANT: Donna T. Ward				
; APPLICANT: Andrew T. Watt				
; TITLE OF INVENTION: ANTISENSE MODULATION OF RECQL2 EXPRESSION				
; FILE REFERENCE: RTS-0207				
; CURRENT APPLICATION NUMBER: US/09/798,096				
; CURRENT FILING DATE: 2001-03-01				
; NUMBER OF SEQ ID NOS: 89				
; SEQ ID NO 10				
; LENGTH: 99500				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURE:				
US-09-798-096-10				
Query Match				
Best Local Similarity 8.2%; Score 245.6; DB 4; Length 99500;				
Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;				
QY	314	GCCGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGCGCGAGGTGGTGGAT	373	
Db	8960	CGCGGGCGTGGTGGCTTATGCTGTAATCCAGCACTTTGGGAGCGCGAGCGGCGAGAT	8901	
QY	374	CACCTGAGTCAAGGATTCAGACACGCTGCGCCACATGCTGAACCTGCTCTACTTA	433	
Db	8900	CACCTGAGTCAAGGATTCAGACACGCTGCGCCACATGCTGAACCTGCTCTCTACTTA	8841	
QY	434	AAAAATAAAATAGCTGGGATGCTGGTGGTACCTGTAATCCAGATACTTTGGGAGG	493	
Db	8840	AAAAATAAAATAGCTGGGATGCTGGTGGTACCTGTAATCCAGATACTTTGGGAGG	8781	
QY	494	TTGAGCAGGAGAATCGCTTGAACCCGGGAGGCGAGATGTCAGTGAAACAAGACTGTGC	553	
Db	8780	CTGAGACATGAGAATCTCTTGAACCCAGCTGGTAGAGTTGCGAGTGAGCGCAAGATCATAC	8721	
QY	554	CACCTGACCTCCAGTCTGGGCAACAGAGTGAAGCTCCATCTCAACAAACAAACAAAGC	613	
Db	8720	CACCTGACCTCCAGCTGGGCGACAGAGTGAGACTCTCTCTCAAAACAAACAAACAA	8664	
QY	614	AGTGCCCATCATGTAGGATTTAGTGTAGTGTAGTGAGGACTGAGCCTTTGTGCAAGTGAGCA	673	
Db	8663	-----CTTCCTCTGGGTAGGACAAAGGGCTGGGAGTGACTTTAATCACCATGGCCAA	8610	
QY	674	CTCACTAATC-ACCAGTTGTAGTATCAGTGATAACCATCAATGATCCAGGTAAAGCCCT	732	
Db	8609	TTATTTAATCAATCATGCTTAGTAAAGGCGCTAGCTAAACCCAGAGGATAGGGCT	8550	
QY	733	GAGGTTTCAGAAAGATCGCGAGCGCTTCAAGGTGCTGGGATTTGTTGGCAAGCCCTC	792	
Db	8549	CAGGGAGCTCTTAGTGTGGTGAGAAATGTAGATGTGGGAGGCTGTGTCCTGGAG	8490	
QY	793	GAATAATAGAAACAGATTCTGTATTACAAACAGAAAGCAGGAGGCCCATGCTGGGTGCTG	852	
Db	8489	AGGGCATGGAAGTGGCCACACCCCTTCCCATACATACCTTGTCTGTATTCTTCCAGCTG	8430	
QY	853	CCAGGAACTCAGTAGTAGTAACCTAAGACAGCAGCGGTGTCTCCCGACGCACTAGGCCA	912	
Db	8429	GCTGTCTTCTGAGTTGTTTATAAATAGCTGATAATCTAGTAAGGAAATGTTCTGTGACTG	8370	
QY	913	GTGGGAAACAGACTCCACACACAGTCCAGCCAGAGTGTGTGAGGCCCAAGATGGGAA	972	
Db	8369	AATGTGAGGACTTACCACTTGTGCTGTAGTCTGAAGTGGGGGTCTGGAATGCGGGTG	8310	
QY	973	GCACGGGGGAAAGGTACAGGTGGGATGGGAGGGGTACAGGCGAAGAGGGTCAGGGCCCA	1032	
Db	8309	GACAGGGAGGACTTAAACATTTGGGGTCTGCACTAACACCTTGGCAGTGTAGTGTGACACCTG	8250	

QY	1033	GGCTGAGGAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACTGACCCCGCATGTTTG	1092	
Db	8249	AGCTAAATGTAGGACACCCAGTGTCTGCTAAGAACTGAACTATTGTGGGAAA-----	8196	
QY	1093	GTGAGGGAGATTTCAGGAAGTCTTCTCGAAGAGAGGCTGTGGAGCTGAGACTCATAGA	1152	
Db	8195	---AAAACACACACATTTGGTGCAGAGAGTGGGAATTTATTATGGGTACAGAGTAG	8139	
QY	1153	TGAGTGGGAGGGTGTTCAGGCGAAGACACGACACCTACAAAAGCATGACTTTGAGAG	1212	
Db	8138	AGAGGAATGGAGTTTCTTTTACAGTGGTATACGCA-----TAGAATG	8096	
QY	1213	AAGCATTCATCCATTCACATGATGAATTTTCAGACTGGGCGACGCTGGCTCATGCTGTAA	1272	
Db	8095	AAATATTGAGCTTTTAAAGAAATAAATTT--GACTAGATGAAGTGGCTCATGCTGTAA	8038	
QY	1273	TCCAGCACTTTGGAAGCTGAATGGGAGAGTACTTTGAGCCTAGGCAATTTGTACAAAG	1332	
Db	8037	TCCAGCACTTTGGGAGCGCAAGGCGAGGATATCGCTTGGAGCCAGGAGTTTGGAGCCAG	7978	
QY	1333	CCTGGGCAACATGGTGGAGCC--TGCCTCCACAAACAAACAAACAAACAAATCAT	1390	
Db	7977	CCTAGGCAACAAAGCAAGACCCAGTATCTACAAAATAAATAAGTTAGCCAGCATGGT	7918	
QY	1391	TATACCTGGTACCATGGGTACCAGGTACATAGAAATGACTCAGGCAGATATGTTGCTTC	1450	
Db	7917	GGTACACTCTCTGCTCTCACCTTCTCGGAGGCTGAGSCAAGGATCATCTGAGCCCA	7858	
QY	1451	TCCTACTGTGGAGAGCGGGCTTATCTGCTAGTAGAACAATAGAGGAGGGAATATAAT	1510	
Db	7857	GGAGTTAGAGGCTGCGAGTGAGCAACGATCATTTCTACTGCACTCCAGCCTGGATAGT	7798	
QY	1511	CCTAAATGAGAGTACAGATTTGAGAGCAACACAGGSCACAGGCATATGTACAGGGGT	1570	
Db	7797	CAAGACCTCTCTGAAAAATAATAAATAACAATAAATTAAGTTAATAAATAAATAA	7738	
QY	1571	AAAGAGGGAATCAGGGAAGGCTTCTCAGAGAAAGGTGACATTTTAAGCC-----	1625	
Db	7737	AAAAAAATCTGCAACATGAGATCACACAGATGAACCTTTGAAGCCATTAAGCAAGTG	7678	
QY	1626	AAGGATGAACGAGTTAGTTTCCACAGGATGGGATGGAAGGGGTGAGATGAGGAGCA	1685	
Db	7677	AAGTAAGCCAGTCACAGGCAAAATACTACATGATCTCCACTGCAACAGAGTATCTAA	7618	
QY	1686	GAGGNACTGCGAGTATAGGCTAGACAGGAGTCTTGAAGCCCTTGAAGGAGTGAGA	1745	
Db	7617	GGCAAACTCATAGAAAGAGAGGGGCAAGTGGTTGCCAGGGTAGGGGCTTTCAGGAAA	7558	
QY	1746	GAAGACCGGCACTGCTAGTGGTTAAGTAACAAAGCTGAGAAAGCCAGGGAATCCCTG	1805	
Db	7557	TGGGGAGTTGTTAATCAAAAAGTATAAAGTTTCAGTAATGCAAAATGAACGATTTCT	7498	
QY	1806	GTCATCGAGGGCTGTGAGTCACTGAGAGTGTTCGGGCTTTTGTCTCTGGGAGCAG	1865	
Db	7497	GATCTGCCATACAACTTTGTGTC--TGTGTGTTACTATACATGTAAAACTCTGTTAA	7439	
QY	1866	TCGATTTTAAAGCAGGGAACAGCTGTATTCAGAGTTTCGGAAGATCTGTGTTGCTGCT	1925	
Db	7438	TTAAGAAACAAATTTTAAAAAATTTTATAGACACAGGTTCTGCTCTGCTCACTCAG	7379	
QY	1926	AAGGGATGAAACTGGAGGCTAGGAGCCAGGGTGATAGGAGGATCCAGGGTGATGGGG	1985	
Db	7378	GCTAGAGCGCAGTGGCACAATCATGTTCACTGAGCCTCAGACCTCTAAGCTCAAGTGA	7319	
QY	1986	AGGCTGGGAGGTCGGGTGATGGACGAG--GCTGGGCGCCAGGGGATGGGAGGAAGG	2042	
Db	7318	TCCTCTGCTCAGCCTCTGGAGAACTGGAATCTGAGGTGCATGACATCATGCTCACT	7259	
QY	2043	AGTAATTTGGGAGAGGCTGGGGCTCTGGCGAGGAATG-----GATGTTGGGTGAAAC	2096	
Db	7258	AAATTTTAAATTTTATAGACAAAGTCTCCCATTTGTTGCCAGACTGGTCTTGAATCTC	7199	

;; TITLE OF INVENTION: ISOLATED HUMAN RECEPTOR PROTEINS,
;; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN RECEPTOR PROTEINS, AND
;; FILE REFERENCE: CL001238
;; CURRENT APPLICATION NUMBER: US/09/851,985
;; CURRENT FILING DATE: 2001-05-10
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 21968
;; TYPE: DNA
;; ORGANISM: Human
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(21968)
;; OTHER INFORMATION: n = A,T,C or G
US-09-851-985-3

Query Match 7.8%; Score 234; DB 4; Length 21968;
Best Local Similarity 82.9%; Pred. No. 2.1e-54;
Matches 267; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 314 GCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCGAGGTGGGTGGAT 373
DB 12119 GCTGGGCGGGTGGCTCAGCCTGTATCCAGACTTTGGGAGGCGAGGTGGAT 12178

QY 374 CACTTGAGTCTAGGAGTTCGAGACCAAGCTGGCCCAACATGGTGAAACCCCTGTCTACTA 433
DB 12179 CAGCTGAGTCTAGGAGTCAAGACCAAGCTGGCCCAACATGGTGAAACCCCTGTCTACTG 12238

QY 434 AATAATATAAATAGTGGGCTAGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 493
DB 12239 AATAATATAAATAGTGGGCTAGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 12298

QY 494 TTGAGCAGGAGATCGCTTGAACCGGAGGCGAGATGTTGAGTGAACCAAGATCTGTGC 553
DB 12299 CTGAGCAGGAGATCACTTGAACCGGAGGCGAGATGTTGAGTGAACCAAGATCACTAC 12358

QY 554 CACTGCCTCCAGTCTGGGCAACAGAGTGAAGCTTCATCTCAACCAACCAACCAACCAAC 613
DB 12359 CACTGCCTCCAGTCTGGGCAACAGAGTGAAGCTTCATCTCAACCAACCAACCAACCAAC 12418

QY 614 AGTGGCCATCATGTAGATTGA 635
DB 12419 ATAGCCAGTTGCCTAGATTGA 12440

RESULT 5
US-09-061-702-1/c
; Sequence 1, Application US/09061702
; Patent No. 6185737
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Liu, Xueqiong
; TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN
; TITLE OF INVENTION: APOPTOSIS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/061,702
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:

;; ATTORNEY/AGENT INFORMATION:
;; NAME: McMillian, Nabeela R.
;; REGISTRATION NUMBER: P-43,363
;; REFERENCE/DOCKET NUMBER: UTSD:546
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (512) 418-3000
;; TELEFAX: (512) 474-7577
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2839 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-09-061-702-1

Query Match 7.8%; Score 232.6; DB 3; Length 2839;
Best Local Similarity 71.9%; Pred. No. 1.7e-54;
Matches 304; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 228 CACCACTTACTGCTGTGACCTTTGCACATGGCTTAGTCCCTCTGTTGCCCTCATCTGTC 287
DB 2361 CAGCAGTTTCAAGTGGTTCACAGTATAGCTGTTCAACAGACATTTGGGGGGTTTCTATT 2302

QY 288 AATGGAGTGATAACAGTGCCCATCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCCAAC 347
DB 2301 TCTTCCCATTAACAATTCAGCAGTTCAGTGCAGTGGCTCATGCTGTATATCCAGC 2242

QY 348 ACTTTGGGAGGCGGAGGTGGTGGATCACTTGAAGTCAAGAGTTCGAGACCAAGCTGGGC 407
DB 2241 ACTTTGGGAGGCGGAGGTGGTGGATCACTTGAAGTCAAGAGTTCGAGACCAAGCTGGGC 2182

QY 408 ACATGCTGAACCCCTGCTCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAA 467
DB 2181 AACATGCTGAACCCCTGCTCTACTATAAATAAATAAATAAATAAATAAATAAATAAATA 2122

QY 468 ACCTGTATATCCAGATATCTTGGGAGTTCGAGGAGGAGTTCGAGGAGGAGTTCGAGGAG 527
DB 2121 GCCTGTATATCCAGATATCTTGGGAGTTCGAGGAGGAGTTCGAGGAGGAGTTCGAGGAG 2062

QY 528 GATGTTGAGTGAACCAAGACTGTGCCACTGCTCCAGTCTGAGGAGTTCGAGGAGTTCGAGG 587
DB 2061 GAGGTTGAGTGAACCAAGACTGTGCCACTGCTCCAGTCTGAGGAGTTCGAGGAGTTCGAGG 2002

QY 588 CCATCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 647
DB 2001 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1942

QY 648 GGA 650
DB 1941 GTA 1939

RESULT 6
US-09-916-204-3
; Sequence 3, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 48763
; TYPE: DNA
; ORGANISM: Human
US-09-916-204-3

Query Match 7.7%; Score 232.2; DB 4; Length 48763;

Best Local Similarity		85.7%;	Pred. No. 1e-53;	
Matches		258;	Conservative	0; Mismatches 43; Indels 0; Gaps 0;
QY	314	GCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACATTTGGGAGGCGGAGGTGGGTGGAT	373	
Db	40920	GCCGGGCGAGTGGCTCACACCTGTAATCCAGCACATTTGGGNGACCGAGGCGGGTGGAT	40979	
QY	374	CAC TTGAGGT CAGAGTTCGAGACCGAGCCTGGGCCAACATGGTGAACCCCTGTCTCTACTA	433	
Db	40980	CAC TTGAGGT CAGAGTTCGAGATGGCCCTGGGCCAACATGAAATCCCGTCTCTACTA	41039	
QY	434	AAAATATAAAATTTAGCTGGCGCATGTGGTGGCGTACTGTAAATCCAGATATCTGGGAGG	493	
Db	41040	AAAATACAAAATTTAGCTGGCGATGTGGCGCATGCTGTAGTCCAGCCATCTCCGAGG	41099	
QY	494	TTGAGGCAGGAGAAATCGCTTTGAACCCGGGAGGCAGATGTTGCAGTGAACCAAGACTGTGC	553	
Db	41100	CTGAGGCAGGAGAAATCGCTTTGAACCCGGGAGGCAGAGTTGCAGTGAACCGAGATCGTGG	41159	
QY	554	CAC TTGCACTCCAGTCTGGGCGACAGAGTGAGCCTCCATCTCAAAACAACAACAAAAAGC	613	
Db	41160	CAC TTGCACTCCAGCTGGGTGAACAGAGTGAGACTTCATCTCAAAAAAACCACAAAAAAC	41219	
QY	614	A 614		
Db	41220	A 41220		

```

RESULT 7
US-09-740-041-3/c
; Sequence 3, Application US/09740041
; Patent No. 6562593
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001001
; CURRENT APPLICATION NUMBER: US/09/740,041
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 66804
; TYPE: DNA
; ORGANISM: Human
US-09-740-041-3

```

Query Match	7.7%;	Score 232;	DB 4;	Length 66804;
Best Local Similarity	85.2%;	Pred. No. 1.4e-53;		
Matches 259;	Conservative 0;	Mismatches 45;	Indels 0;	Gaps 0;

QY	314	GC	CGGGTGCAGTGCCTAGTGCCTGAAATCCCA	CACTTTGGAGGCGGAGGTGGGTGGAT	373
Db	31625	GT	GGCGCAGTGGCTCATGCTCTGAATCCCGACATTTGGGAGCCCAAGCGGGTAGAT		
QY	374	CAC	TTGAGTTCAGGAGTTTCGAGACCAGCAGCTGC	CCAACTGCTGTAACCCCTGTCTCTACTA	433
Db	31565	CAC	TTGAGTTCAGGAGTTTCGAGACCAGCAGCTGC	CCAACTGCTGTAACCCCTGTCTCTACTA	
QY	434	AAA	ATATAAAAATTAGCTGGGCACTGGTGGTGGTACCTGT	TAATCCCGATACTTTGGCAGG	493
Db	31505	AAA	ATATAAAAATTAGCTGGGCACTGGTGGC	CATGCTGTAACTCCAACTACTCAGAGG	31446
QY	494	TT	GAGCAGGAGAAATCGTTTGAACCCGGAGCGCAGATGTT	CGACTGAAACCAAGACTGTGC	553
Db	31445	CT	GGGCGAGGAAATCGTTTGAACCTGGGAGCGGAGTTC	CGACTGAGCCAAAGATTGTGC	31386
QY	554	CAC	TGCATCCAGTCTGGGCACAGATGAGCCTTCATCT	CAACAAACAAACAAACG	613
Db	31385	CAC	TGCATCCAGCCTGGGTGACAGATGAAGACCCCATCT	CAAAAAAAGGAAA	31326
QY	614	AGT	G 617		

Db 31325 TGTG 31322

```

RESULT 8
US-09-009-913-1/c
; Sequence 1, Application US/09009913
; Patent No. 6087485
; GENERAL INFORMATION:
; APPLICANT: Axys Pharmaceuticals, Inc.
; TITLE OF INVENTION: Asthma Related Genes
; NUMBER OF SEQUENCES: 339
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,913
; FILING DATE: 21-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-4P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3231
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72928 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-009-913-1

```

Query Match	7.7%;	Score 231.4;	DB 3;	Length 72928;
Best Local Similarity	85.2%;	Pred. No. 2.2e-53;		
Matches 270;	Conservative 0;	Mismatches 46;	Indels 1;	Gaps 1;
QY	296	TCGATAACAGTGC	CCCATACACCGGGTGCAGTGCCTAGTGCCTGA	AATATCCCAACACCTTTGGG 355
Db	38968	TAAATAAGATGCTCAATCGGCCAGGTGCAGTGGCTCGTCTGTAATCCAGACACTTTGGG		
QY	356	AGGCGGAGGTGGTGGATCACTTTGAGGTCAGGAGTTCGAGACCAAGCTCGGCCAAATGCT 415		
Db	38908	AGGCCGAGCGAGCGAGATCACTTTGAGGTCAGGAGTTTGAGACCAAGCTTCACCAACATGCT 38849		
QY	416	GAACCCCTGCTCTACTAAAAATATAAAATTTAGCTGGGCATGGTGGTGCCTGACCTGTAA 475		
Db	38848	GAACCCCTGCTCTACTAAAAATATAAACTTTAGCCAGGCGATGGTGGTGCATGCTGTAA 38789		
QY	476	TCCCGATACATTGGGAGGTTGAGGCAGGAGAAATCGCTTTGAACCCCGGGAGGCAGATGTTGC 535		
Db	38788	TCCCATCTACTTTGGGAGGCTGAAGCAGGAGAAATCACTTTGAACCCCGGGAGGCGGAGTTGC 38729		
QY	536	AGTGAACCAAGACTGTGCGCACTGCACCTCCAGTCTGGGCAAC-AGAGTGAGGCTCCATCTC 594		
Db	38728	AGTGAGCCAAGATAGCACCATTTGCACCTCCAGCTGGGCAACAGAGCAAACTCCATCTC 38669		
QY	595	AAACAAACAAACAAAAA 611		

Db 10992 GAGCGGGTGGATCAGGAGTTCAGGATCGAGATCGAGACCATCTCGCCAAACATGTGTAA 11047

RESULT 14

US-08-724-394A-20/c

; Sequence 20, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H. CONTIG"

US-08-724-394A-20

Query Match 7.6%; Score 229; DB 2; Length 246240;

Best Local Similarity 82.6%; Pred. No. 2e-52;

Matches 262; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 295 GTGATAACAGTGGCCATCAGCCGGTGCAGTGGCTAGTGGCTGAAATCCCAACATTTGG 354

Db 2941 GTATTAGGATGAACCCCGCTGGCGCGTGGCTCACCTGTAATCCAGCACTTTGG 2882

Qy 355 GAGCGGAGTGGTGGATCACTTGGATCGAGGTTTCGAGACCATCTCGCCAAACATGG 414

Db 2881 GAGCGGAGGCGGTAGATCACTTGGATCGAGGTTTCGAGACCATCTCGCCAAACATGG 2822

Qy 415 TGAACCTGTCTCTACTATAAAATATAAAATAGCTGGGATCGTGGTACCTGTGA 474

Db 2821 CGAAACCTGTCTCTACTATAAAATATAAAATAGCTGGGATCGTGGTACCTGTGA 2762

Qy 475 ATCCAGATATTCGGAGGTTGAGCGAGGAGATTCGCTTGAACCCGGGAGGCAATGTTG 534

Db 2761 ATCCAGCTGCTTGGGAGCTGAGCGAGGAGATTCGTTTGAACCTGGACGCTGGGATTTG 2702

Qy 535 CAGTGAACCAAGACTGTGCCACTGCACCTCAGTCTGGGCAACAGAGTGAGCCTTCCATCTC 594

Db 2701 CAGTGAGCCAAAGACTGTGCCACTGCACCTCAGTCTGGGCGACAGGGTGAGACAACTCTC 2642

Qy 595 AAACAACAAACAAACAAA 611

Db 2641 AAAAAATAAAAAATAAA 2625

RESULT 15

US-08-724-394A-21/c

; Sequence 21, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H. CONTIG"

US-08-724-394A-21

Query Match 7.6%; Score 229; DB 2; Length 246240;

Best Local Similarity 82.6%; Pred. No. 2e-52;

Matches 262; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 295 GTGATAACAGTGGCCATCAGCCGGTGCAGTGGCTAGTGGCTGAAATCCCAACATTTGG 354

Db 2941 GTATTAGGATGAACCCCGCTGGCGCGTGGCTCACCTGTAATCCAGCACTTTGG 2882

Qy 355 GAGCGGAGTGGTGGATCACTTGGATCGAGGTTTCGAGACCATCTCGCCAAACATGG 414

Db 2881 GAGCGGAGGCGGTAGATCACTTGGATCGAGGTTTCGAGACCATCTCGCCAAACATGG 2822

Qy 415 TGAACCTGTCTCTACTATAAAATATAAAATAGCTGGGATCGTGGTACCTGTGA 474

Db 2821 CGAAACCTGTCTCTACTATAAAATATAAAATAGCTGGGATCGTGGTACCTGTGA 2762

Qy 475 ATCCAGATATTCGGAGGTTGAGCGAGGAGATTCGCTTGAACCCGGGAGGCAATGTTG 534

Db 2761 ATCCAGCTGCTTGGGAGCTGAGCGAGGAGATTCGTTTGAACCTGGACGCTGGGATTTG 2702

Qy	415	TGAAACCCCTGCTCTACTATAAAATATAAAATTAGCTGGGCATGGTGGTACCTGTA	474
Db	2821	CGAAACCCCTGCTCTACTATAAAATACAAAATTAGCTGGGCATGGTGGCAGCACCTGTA	2762
Qy	475	ATCCAGATACTTGGAGGTTGAGGCAGGAGAAATCGCTTGAAACCCGGGAGGCAGATGTTG	534
Db	2761	ATCCAGCTGCTTGGGAGACTGAGGCAGGAGAAATCGTTTGAACCTGGGACGTTGGGATTG	2702
Qy	535	CAGTGAACCAAGACTGTGCCACTGCCACTCCAGTCTGGGCAACAGAGTGAGCCTCCATCTC	594
Db	2701	CAGTGAGCCAGACTGCACCACTGCCTCCAGCCTGGGCGACAGGGTGAGACAACTCTC	2642
Qy	595	AAACAAACAAACAAAA	611
Db	2641	AAAAATAAAAATAAA	2625

Search completed: September 17, 2004, 19:04:56
 Job time : 202.326 secs

This Page Blank (uspto)

1. *Introduction*

QY	121	AGGTCAGGCTGAGGAGTTTCAGCGCCCGGGAAACCGTAGCGATGCTGGGAAGGAGCTTTTGA	180	Db	3180	TGACTTTGAGAGAAGCAATTTCCTCATTCAACTGATGAATTTTCAGACTGGGACCGCTGC	3239
Db	2100	AGGTGAGGCTGAGGAGTTTCAGCGCCCGGGAAACCGTAGCGATGCTGGGAAGGAGCTTTTGA	2159	QY	1261	TCATGCCCTGTAATCCCAAGCACTTTTGGAAAGCTGTAATGGGGAGGATGACTTGTAGCCCTAGGC	1320
QY	181	TGGCCATGTTAGTCAAGGGCTGTAGGCCCTCCGCTCACAGCTGCCACCACTTACTGG	240	Db	3240	TCATGCCCTGTAATCCCAAGCACTTTTGGAAAGCTGTAATGGGGAGGATGACTTGTAGCCCTAGGC	3299
Db	2160	TGGCCATGTTAGTCAAGGGCTGTAGGCCCTCCGCTCACAGCTGCCACCACTTACTGG	2219	QY	1321	ATTGTGTGACAGCCCTGGGCAACATGTTGAGACCTTCCCTCCACAAAACAAAACAAAC	1380
QY	241	TGTGTGACCTTTGACATGGCTTAGTCCCTCTGTGGCTCATCTGTCAAATGGAGTGATA	300	Db	3300	ATTGTGTGACAGCCCTGGGCAACATGTTGAGACCTTCCCTCCACAAAACAAAACAAAC	3359
Db	2220	TGTGTGACCTTTGACATGGCTTAGTCCCTCTGTGGCTCATCTGTCAAATGGAGTGATA	2279	QY	1381	AAAAAATCATTTATCTCTGTTACCATGGGTACCAAGGTACATAGAAATGACTCAGGCAGATA	1440
QY	301	ACAGTGCCTCATCAGCCGGGTGCAAGTGGCTTAGTGGCTTGAATCCCAACACTTTGGGAGGCG	360	Db	3360	AAAAAATCATTTATCTCTGTTACCATGGGTACCAAGGTACATAGAAATGACTCAGGCAGATA	3419
Db	2280	ACAGTGCCTCATCAGCCGGGTGCAAGTGGCTTAGTGGCTTGAATCCCAACACTTTGGGAGGCG	2339	QY	1441	TGGTGTCTCTCTACTGTGTGAGAGGCGGGCTTACTTGCAGTAAGACAATAGAGGGAG	1500
QY	361	GAGTGGGTGGATCACTTCAGCTCAGGAGTTCGAGACCAAGCTGGCCCAACATGGTGAAC	420	Db	3420	TGGTGTCTCTCTACTGTGTGAGAGGCGGGCTTACTTGCAGTAAGACAATAGAGGGAG	3479
Db	2340	GAGTGGGTGGATCACTTCAGCTCAGGAGTTCGAGACCAAGCTGGCCCAACATGGTGAAC	2399	QY	1501	GGAAATATATCTCTAAATGAGAGGTACAGATTTGAGAGCAAAACACAGGGCACAGGCATAT	1560
QY	421	CCTGTCTCTACTAAATAATAAAATTTAGCTGGGCAATGGTGGCTTACTGTAAATCCCA	480	Db	3480	GGAAATATATCTCTAAATGAGAGGTACAGATTTGAGAGCAAAACACAGGGCACAGGCATAT	3539
Db	2400	CCTGTCTCTACTAAATAATAAAATTTAGCTGGGCAATGGTGGCTTACTGTAAATCCCA	2459	QY	1561	GTACGAGGGTAAAGAGGGAAATCAGGGAAGCTTCTCAGAGAAGGTGACATTTAAGCCGGG	1620
QY	481	GATACTTGGGAGGTTGAGGAGGAGAAATCGCTTGAACCCGGGAGGCGAGATGTTGCAAGTGA	540	Db	3540	GTACGAGGGTAAAGAGGGAAATCAGGGAAGCTTCTCAGAGAAGGTGACATTTAAGCCGGG	3599
Db	2460	GATACTTGGGAGGTTGAGGAGGAGAAATCGCTTGAACCCGGGAGGCGAGATGTTGCAAGTGA	2519	QY	1621	ACATGAAGGATGAACAGATTTAGTTTACCAAGGATGGATGGAAAGGGGTGAGAGTGATGG	1680
QY	541	ACCAAGACTGTGCCACTGCACCTCAGCTGTGGCAACAGAGTGAGCTTCCATCTCAAACAA	600	Db	3600	ACATGAAGGATGAACAGATTTAGTTTACCAAGGATGGATGGAAAGGGGTGAGAGTGATGG	3659
Db	2520	ACCAAGACTGTGCCACTGCACCTCAGCTGTGGCAACAGAGTGAGCTTCCATCTCAAACAA	2579	QY	1681	AGGCACAGGGAACCTGCAGGATCATAGGCCCTAGACAGGGATCTCTGACGCCCTTTGAGGAAG	1740
QY	601	ACAAAACAAAAGCAGTGCCTCATCATGTAGGATTTAGTGTAGTGAGTGAAGGACTGAGCCTTG	660	Db	3660	AGGCACAGGGAACCTGCAGGATCATAGGCCCTAGACAGGGATCTCTGACGCCCTTTGAGGAAG	3719
Db	2580	ACAAAACAAAAGCAGTGCCTCATCATGTAGGATTTAGTGTAGTGAGTGAAGGACTGAGCCTTG	2639	QY	1741	TGAGAGAAGACCAAGCCAGCTGTAAGTGAAGTAAACAAAGCTGAGAAAGCCAGGGAAT	1800
QY	661	TGCAAAAGTGAGACTCATTAATCACAGGTTGTAGTATCAGTATTAACCATCAATGATCC	720	Db	3720	TGAGAGAAGACCAAGCCAGCTGTAAGTGAAGTAAACAAAGCTGAGAAAGCCAGGGAAT	3779
Db	2640	TGCAAAAGTGAGACTCATTAATCACAGGTTGTAGTATCAGTATTAACCATCAATGATCC	2699	QY	1801	CCCTGGTCTATGACAGGCCCTGTGATCACGTTCAGGTGTTTGGCTTTTGTCTTCTGGG	1860
QY	721	AGTTAAAGCCCTCAGGGTTTCAGAAAGATGCCGAGCGCTTTCAAGGTGCTTGGGGATTGGT	780	Db	3780	CCCTGGTCTATGACAGGCCCTGTGATCACGTTCAGGTGTTTGGCTTTTGTCTTCTGGG	3839
Db	2700	AGTTAAAGCCCTCAGGGTTTCAGAAAGATGCCGAGCGCTTTCAAGGTGCTTGGGGATTGGT	2759	QY	1861	AGCAGTCTGATTTTAAAGCAGGGAACAGCTGTATTTACAGATTGGGAAGATCCTGTGTTGCT	1920
QY	781	GGGCAAGCCCTCGAATAATAGAAACAGTTCTCTGTATTACAAAGAAAGCAGAGGCCCA	840	Db	3840	AGCAGTCTGATTTTAAAGCAGGGAACAGCTGTATTTACAGATTGGGAAGATCCTGTGTTGCT	3899
Db	2760	GGGCAAGCCCTCGAATAATAGAAACAGTTCTCTGTATTACAAAGAAAGCAGAGGCCCA	2819	QY	1921	GCCTGAAGGGGATGAAAATGAGGCTTAGGAGCCAGGCTCATAGGGAGGATCCAGGGTGA	1980
QY	841	TGCTGGGTGCTGCCAGGAACTCAGTAGTAATTAAGACAGCAGCAGCGGTCTGTTCCCCAGC	900	Db	3900	GCCTGAAGGGGATGAAAATGAGGCTTAGGAGCCAGGCTCATAGGGAGGATCCAGGGTGA	3959
Db	2820	TGCTGGGTGCTGCCAGGAACTCAGTAGTAATTAAGACAGCAGCAGCGGTCTGTTCCCCAGC	2879	QY	1981	TGGGGAGGCTGGGAGGTCGCGGTGATGGACACAGGCTGGGGCCAGGGGATGGGAGGAA	2040
QY	901	GCACCTAGGCCAGTGGGGAACAGACTCACACACAGTCCAGCCACAGAGTGTGAGGGC	960	Db	3960	TGGGGAGGCTGGGAGGTCGCGGTGATGGACACAGGCTGGGGCCAGGGGATGGGAGGAA	4019
Db	2880	GCACCTAGGCCAGTGGGGAACAGACTCACACACAGTCCAGCCACAGAGTGTGAGGGC	2939	QY	2041	GGAGTAATTCGGGAGAGCCCTCTGGGCTCTGGCCGAGGAATGATGGTGGGCTGAAACAGG	2100
QY	961	CAAGATCGGGAAGCACCGGGAGAAAGTTCAGGTTGGGATGGGAGGGGTGAGGGCAAGAG	1020	Db	4020	GGAGTAATTCGGGAGAGCCCTCTGGGCTCTGGCCGAGGAATGATGGTGGGCTGAAACAGG	4079
Db	2940	CAAGATCGGGAAGCACCGGGAGAAAGTTCAGGTTGGGATGGGAGGGGTGAGGGCAAGAG	2999	QY	2101	AGAGGAGAGATGCTTAGGCCACTTTTGGAAACACAGTAGGGCAGGACACAGGAGACCCCAAG	2160
QY	1021	GGGTGAGGCCAGACTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGAC	1080	Db	4080	AGAGGAGAGATGCTTAGGCCACTTTTGGAAACACAGTAGGGCAGGACACAGGAGACCCCAAG	4139
Db	3000	GGGTGAGGCCAGACTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGAC	3059	QY	2161	GGGAGGTGCCCAAGAGACCCAGCAGGCTGCGATTTGGACAGGGAGGTCGTCTGTGGAGCA	2220
QY	1081	CCGGCATGTTTGGTGAGGGAGATTTCAGGAAGTCTTCTTGGGAAGAGAGGCTGTGAGAGCTG	1140	Db	4140	GGGAGGTGCCCAAGAGACCCAGCAGGCTGCGATTTGGACAGGGAGGTCGTCTGTGGAGCA	4199
Db	3060	CCGGCATGTTTGGTGAGGGAGATTTCAGGAAGTCTTCTTGGGAAGAGAGGCTGTGAGAGCTG	3119	QY	2221	GGTGTCTTGGATAAGGGAGGAAAATGGTCAGTTTCCATTCCTCCTCTCTGTGTCAACC	2280
QY	1141	AGACTCAATAGATGAGTGGGAGGGTGTTCAGGCAGAAAGACACAGCACCTTACAAAGCA	1200	Db	4200	GGTGTCTTGGATAAGGGAGGAAAATGGTCAGTTTCCATTCCTCCTCCTCTCTGTGTCAACC	4259
Db	3120	AGACTCAATAGATGAGTGGGAGGGTGTTCAGGCAGAAAGACACAGCACCTTACAAAGCA	3179	QY	2281	TCTAAACTACATGGGGCACAGGACCCAGTGGGACTCCATTAATGATGGGATGGGTGGATG	2340
QY	1201	TGACTTTGAGAGAGCAATTCATCCATTCAACTGATGAATTTTTCAGACTGGGCAAGCTGGC	1260				

D _b	4260	TCTAAACTACATGGGCGACAGGACCCAGTGGGAATCCTCAATAAATGATGGGATGGGTGGATG	4319
Q _y	2341	GAAGGAAGGAAGAGAGAAACAACCTCTTCATTATCATCTGGTTATTTACAGAACAGGCCAGG	2400
D _b	4320	GAAGGAAGGAAGAGAGAAACAACCTCTTCATTATCATCTGGTTATTTACAGAACAGGCCAGG	4379
Q _y	2401	TGC GG T G C T C A C G C T T G C C A T T C T A G C A C T T T T G G A G G C T G A G G T G G A T T A C C T C A	2460
D _b	4380	TGC GG T G C T C A C G C T T G C C A T T C T A G C A C T T T G G A G G C T G A G G T G G A T T A C C T C A	4439
Q _y	2461	G G T C A G A G T T C A G A C C A G C C T I A G A C A A C G T A G A G A A A C C C A T C T C T A C T G A A G A T A T	2520
D _b	4440	G G T C A G A G A T T C A A G C C A G C C T A G A C A A C G T A G A G A A A C C C C A T C T C T A C T G A A G A T A T	4499
Q _y	2521	A A A A T T A G C T T G G G C G T A G T G C C A T A T C C C T G T A A T C C C A G C T A G T C G G G A A G C T G A G C C A	2580
D _b	4500	A A A A T T A G C T T G G G C G T A G T G C C A T A T C C C T G T A A T C C C A G C T A G T C G G G A A G C T G A G C C A	4559
Q _y	2581	G G A G A A T C G C T T G A A C C C G A G A G C A G A G G T T G C G G T G A G C T G A G A T C G T G C C A T T G C A C	2640
D _b	4560	G G A G A A T C G C T T G A A C C C G A G A G C A G A G G T T G C G G T G A G C T G A G A T C G T G C C A T T G C A C	4619
Q _y	2641	T C C A G C C T G G G T G A C A A A G C A A G A C C T C G T C T C A A T A A T A A T A A T A A T A A A A A C A G A A	2700
D _b	4620	T C C A G C C T G G G T G A C A A A G C A A G A C C T C G T C T C A A T A A T A A T A A T A A T A A A A A A C A G A A	4679
Q _y	2701	G G A G C C T G G G T C A T C C C A G C T A C C T A C T T T T C A G G A G A A T G T A C T C C C T T A C C A C C A G G C C	2760
D _b	4680	G G A G C C T G G G T C A T C C C A G C T A C C T A C T T T T C A G G A G A A T G T A C T C C C T T A C C A C C A G G C C	4739
Q _y	2761	A A A G A T G G G A G A A C C A G T T T G A T T A T G C A T T T A T T G A G C A C C T A C T T G A G T C C C A T C C C	2820
D _b	4740	A A A G A T G G G A G A A C C A G T T T G A T T A T G C A T T T A T T G A G C A C C T A C T T G A G T C C C A T C C C	4799
Q _y	2821	T G G G C T A G G C T G G A A T G G A C T C A G A T G G A G C C T G A A G A G T C C C C C T C A G G A A C C T C A C T	2880
D _b	4800	T G G G C T A G G C T G G A A T G G A C T C A G A T G G A G C C T G A A G A G T C C C C C T C A G G A A C C T C A C T	4859
Q _y	2881	A G A A A A G A G A G A A T C G G C C G G C G G T G G C T C A G C C T G T A A T C C C A A C T T T T G G G	2940
D _b	4860	A G A A A A G A G A G A A T C G G C C G G C G G T G G C T C A G C C T G T A A T C C C A A C T T T T G G G	4919
Q _y	2941	A G G C T A G G T G G G T G G A T C A A A G G T C A G A G A T C G A G A C C A T C C T G G C T T A C A C A G T G A	3000
D _b	4920	A G G C T A G G T G G G T G G A T C A C A A G T C A G A G A T C G A G A C C A T C C T G G C T T A C A C A G T G A	4979
Q _y	3001	A 3001	
D _b	4980	A 4980	

	RESULT 2	
	US-10-669-693-3	
	; Sequence 3, Application US/10669693	
	; Publication No. US20040043413A1	
	; GENERAL INFORMATION:	
	; APPLICANT: YAN, Chunhua et al	
	; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING	
	; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN	
	; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,	
	; TITLE OF INVENTION: AND USES THEREOF	
	; FILE REFERENCE: CL000685CON	
	; CURRENT APPLICATION NUMBER: US/10/669,693	
	; CURRENT FILING DATE: 2003-09-25	
	; PRIOR APPLICATION NUMBER: 09/748,127	
	; PRIOR FILING DATE: 2000-12-27	
	; NUMBER OF SEQ ID NOS: 4	
	; SOFTWARE: FastSeq for Windows Version 4.0	
	; SEQ ID NO 3	
	; LENGTH: 17752	
	; TYPE: DNA	
	; ORGANISM: Homo sapiens	
	; FEATURE:	
Db	2580	ACAAACA...AGCAGTGCCTCATGTAGGATTGATTTAGTGAGGACTGAGCCCTTG 2639
QY	661	TGCAAGTGACCACTCACTAATCACCGAGTTGTAGTATCACTGATAACCATCAATGATCC 720
Db	2640	TGCAAGTGACCACTCACTAATCACCGAGTTGTAGTATCACTGATAAACCACTCAATGATCC 2699
QY	721	AGGTAAAGCCCTGAGGGTTTCAGAAAAGATGCCGGAGCGCTTTTCAAGGTGCTCGGGATTGGT 780
Db	2700	AGGTAAAGCCCTGAGGGTTTCAGAAAAGATGCCGGAGCGCTTTTCAAGGTGCTCGGGATTGGT 2759
QY	781	GGGCAAGCCCTCGAATAATAGAAAACAGTTCTCTATTACAAACAGAAGCAGGAGGCCCA 840
Db	2760	GGGCAAGCCCTCGAATAATAGAAAACAGTTCTCTATTACAAACAGAAGCAGGAGGCCCA 2819
QY	841	TGCTGGGTGCTGCCAGGAACCTCAGTAGTAATAAGACAGACCCGGTCTGTTCTCCAGC 900
Db	2820	TGCTGGGTGCTGCCAGGAACCTCAGTAGTAATAAGACAGACCCGGTCTGTTCTCCAGC 2879
QY	901	GCACCTAGGCCAGTGGGGMAACAGACTCACACACAGTCCCAGCCCCAGAGTGGTCAAGGC 960
Db	2880	GCACCTAGGCCAGTGGGGMAACAGACTCACACACAGTCCCAGCCCCAGAGTGGTCAAGGC 2939


```

; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 61103
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(61103)
; OTHER INFORMATION: n = A,T,C or G
; US-10-087-192-58

```

	Query Match	9.4%;	Score 282;	DB 13;	Length 61103;
	Best Local Similarity	19.7%;	Pred. No. 5.1e-74;	Mismatches 1835;	Gaps 4;
	Matches 457;	Conservative 0;			
QY	387	GAGTTCGAGACCACTGGCCACATGGTGAAACCCTGTCTACTATAAAA-TATAAAAA	445		
Db	41155	GAGATCGAGACCATCTGGACAAATGATGAGTAACCTCCGTATCTACTAAAGAATATAAAA	41096		
QY	446	TTAGCTGGGCATGGTGGTGCCTACCTGTAAATCCAGATACTTTGGGAGGTTGAGGCAGAG	505		
Db	41095	TTAGCTGGGCATGGTGGCATGCATCTGTAGTCCAGCTACTTAGAGGCTGAGGCAGAG	41036		
QY	506	AATCGTTGAAACCCGGAGGCAGATGTTCCAGTGAACCAAGACTGTGCCACTGCACCTCCA	565		
Db	411035	AATTGTTCAAACCCAGGAGCAGAGGTTCAAGTGTAGCCCAAGATCGGCCACTGCAANNNN	40976		
QY	566	GTCGGGCACAGATGAGCTCCTCATCTCAAACAACAACAAAGACGATGCCCATCAT	625		
Db	40975	NN	40916		
QY	626	GTAGGATTGAGTGATTGAGTGAGGACTGAGCCCTTGTCGAAAGTGAGCACTCACTAATCAC	685		
Db	40915	NN	40856		
QY	686	CAGGTTGTAGTATCACTGATGAACCATCAATGATCCAGGTAAAGCCTCAGGGTTCAGAAA	745		
Db	40855	NN	40796		
QY	746	GATCGCGGAGCGCTTCAAGGTGTGGGGATTGGTGGGCAAGCCCTCGAAATAATAGAAAC	805		
Db	40795	NN	40736		
QY	806	AGTCTCTGTATTACAACAGAAAGCAGAGGCGCCATGCTGGGTGCTGCCAGGAJCTCAGT	865		
Db	40735	NN	40676		
QY	866	AGTAATAAGACAGCACCGTGTCTCTCCCAGCGCACCTAGGCCAGTGGGGAAAACAGA	925		
Db	40675	NN	40616		
QY	926	CTCACACACAGTCCCGACCCAGAGTGGTTCAGGGCCAAAGATGGGGAAGCACGGGAGAAA	985		
Db	40615	NN	40556		
QY	986	GGTCAGGTGGGATGGGAGGGGTCAAGGCAGAGGGGTCAAGGCCAGGCTGAGGGGAAGC	1045		
Db	40555	NN	40496		
QY	1046	CCTGGCACTGTAGGAATTTAGAGGAGGTACCTGACCGGCATGTTTGGTGGAGGAGATTC	1105		

[illegible]


```
;/ NAME/KEY: SITE
;/ LOCATION: (14067)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14068)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14069)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14070)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14071)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14072)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14073)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14074)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14075)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14076)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14077)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14078)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14079)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14080)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14081)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14082)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14083)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14084)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14085)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14086)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14087)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14088)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14089)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14090)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
```

```
;/ LOCATION: (14091)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14092)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14093)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14094)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14095)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14096)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
;/ LOCATION: (14097)
;/ OTHER INFORMATION: n equals a,t,g, or c
;/ NAME/KEY: SITE
```

```
Query Match      8.4%; Score 253; DB 9; Length 16747;
Best Local Similarity 27.2%; Pred. No. 1.9e-65;
Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;

Qy 301 ACAGTGGCCATCAGCGGGTGCAGTGGCTAGTGGCTGAAATCCCAACACATTTGGGAGGCG 360
Db 16152 AAAATGAGTTTAGGCTGGGCGGCTGCTCAGGCTATATCCCAACATTTGGGAGGCC 16093

Qy 361 GAGTGGGTGGATCATTGAGTTCAGAGTTCAGAGCAGCCTGGCCAAACATGTGAAAC 420
Db 16092 GAGCTGGTGGATCATTGAGTTCAGAGTTCAGAGCAGCCTGGCCAAACATGTGAAAC 16033

Qy 421 CCTGCTCTACTAAAAATA-TAAAAATTAGTGGGCATGTGGTGCCTACCTGTAAATCCC 479
Db 16032 TCTGCTCTACTAAATTTACAAAAATTAGTGGGCATGTGGTGGCGGTGCTGTAAATCCC 15973

Qy 480 AGATACTTGGGAGGTTGAGGCAGGAGATCGCTTGAACCCGGGAGGCAGATGTTGCAGTG 539
Db 15972 AGCTACTTGGGAGGCTGAGGCAGAGAAATTCCTGTAACCCAGAGATGGAGGTTGTAGTG 15913

Qy 540 AACCAAGACTGTGCCACTGCCTCCAGTCTGGGCAACA-GAGTGAGCCTCCATCTCAAAAC 598
Db 15912 AGCCAAGACTGCGCCACTGCCTCTAGCCTGGGCAACAACGAGCGGAACTCCGCTCAAAA 15853

Qy 599 AAACAAACAAAAGACAGTGCCTCATGTAGGATTTGAGTGTGATTGAGTACGAGACTGAGCCT 658
Db 15852 AAAAAAATAAATTTGAGGAGATGAGATTTAAATGTAATATATATAGACATCATTTTA 15793

Qy 659 TGTGCAAGTGCAGCTCACTAATCACCAGGTTGTAGTATCAGTGATAACCATCAATGAT 718
Db 15792 TAGGAAATAGTGAGTCTCTGAAGACCTCTATGATGGGTTAATTTGGGGAAGATGTTCAATT 15733

Qy 719 CCAGGTAAGCCCTGAGGTTTCAGAAAGATGCGGAGCGCTTTCAAGGTGCTGGGGATTG 778
Db 15732 TCTGAGTCAAAAAAATTTTACAGACAAATTTTCTGTCTCAGTTCAGAGAGTGGAGAACAGA 15673

Qy 779 GTGGCAAGCCCTCGAATAATAGAAACAGTTCTCTATTATACACAGAAAGAGGAGGCC 838
Db 15672 ATGAGGCAGAAAGGCGATTGAGAGGTGTGCATTATTTTCAGGTGAGAGGTAAGACAG 15613

Qy 839 CATGCTGGGTGCTGCCAGGAACCTCAGTAGTAACTAAGACAGACACCGGTGCTGCTTCCCCA 898
Db 15612 TATCAGTAGAGGTGAAGAGGTTAAGGACACATGCCAAGATTAAATCAAGATTAATA 15553

Qy 899 GCGCACCCTAGGCCAGTGGGGAAACAGACTCAACACAGTCCACAGCCAGAGTGTGAGG 958
Db 15552 TTGAAGGTAATTGAGGTTTTTGTCTCGGGGCACGCCACTAGTTAAGTAATATATAGTTGACA 15493

Qy 959 GCCAAGATCGGAGACACGGGGAG-----AAAGGTGAGGTGGATGGGGAGGG 1008
Db 15492 GGCAGTTTGGAGATAAGATCTGTTGAATTTCTTGTCTGTAAAGGCTTTCTGAGGATTC 15433
```



```
; NAME/KEY: misc_feature
; LOCATION: (14071)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14072)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14073)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14074)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14075)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14076)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14077)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:

Query Match      8.4%; Score 253; DB 16; Length 16747;
Best Local Similarity 27.2%; Pred. No. 1.9e-65;
Matches 659; Conservative 0; Mismatches 1740; Indels 22; Gaps 6;

QY 301 ACAGTGGCCCATCAGCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTTGGGAGGCG 360
DB 16152 AAAATGAGTTAGGCTGGCGGGTGGCTCAGGCTTAAATCCCAACACTTTTGGGAGGCG 16093

QY 361 GAGGTGGGTGATCATCTAGTTCAGGTTCAGACACAGCTGGCCAAACATGGTGAAC 420
DB 16092 GAGGCTGGTGGATCATCTAGGTTCAGGAGTTCAGACACAGCTGGCCAAACATGGTGAAC 16033

QY 421 CCGTCTCTACTAAAAATA-TAAAAATTAGCTGGGCATGGTGGTACCTGTATATCCC 479
DB 16032 TCTGTCCTCTTAAATTTACAAAAATTAGCTGGGCATGGTGGCGGTGCTGTATATCCC 15973

QY 480 AGATACTTGGGAGGTTGAGCGAGGAGATCGCTTGAACCCGGGAGGAGCATGTTGCAGTG 539
DB 15972 AGCTACTTGGGAGGCTGAGCGAGAGAAATTGCTGAACCCAGAGATGGAGGTTGTAGTG 15913

QY 540 AACCAAGACTGTGCCACTGCCTCAGCTTCGGCAACA-GAGTGAAGCTCCATCTCAAC 598
DB 15912 AGCCAGACTGGCCCACTGCCTAGCTTGGGCAACAAGAGCGGAACCTCCGCTCAAAA 15853

QY 599 AAACAAACAAAAGCAGTGCCTCATGTAGGATTGAGTGATTGAGTGAAGGAGTGAAGCT 658
DB 15852 AAAAAAATAAATTCGGGAGATCAGATGTAAATGTAAATATATAGACATCATTTTA 15793

QY 659 TGTGAAAGTGACATCTACTAATACCAAGTTGTAGTATCAGTGATGATAACCATCAATGAT 718
DB 15792 TAGGAAATAGTGAAGTCTCTGAAGACCTCTATGATGGGTAAATGGGGAAGATGTTCAAT 15733

QY 719 CCAGGTAAAGCCCTGAGGTTTCAGAAAGATGCGGAGCGCTTCAAGGTGCTGGGATTG 778
DB 15732 TCTAGTCAAGAAAAATTTTACAGACAAATTTTCTGTCTCAGTTCAGAGGATGGAGAACAG 15673

QY 779 GTGGGCAAGCCCTCGAATAATAAGAACAGTTCTCTGTATTACAAACAGAAAGCAGGAGGC 838
DB 15672 ATGGAGCGAAGAGGCAATTGAGAGGTTGGTGCATTAATTCAGGTGAGAGGTAAAGACAG 15613

QY 839 CATGCTGGGTGCTGCCAGGAACCTCAGTGTAGTAACTAAGACAGACACCGGTGCTGCCCA 898
DB 15612 TATCAGTAGAGGTGAAAGAGGTAAAGGACACATGCCAAGATTAAATCAAGGATAAATA 15553

899 GGCACCTAGCCAGTCAGTGGGAAAACAGACTCACACACAGTCCAGCCAGAGTGTGTGAGG 958
15552 TTGAAGGTAAATTGAGGTTTTTCTCGGGCAGCCCACTAGTTAAGTAAATATATAGTTGGACA 15493
959 GCCAAGATCGGGAAGCAGCGGGAG-----AAAGTCAAGGTGGGATGGGAGGGG 1008
15492 GGCAGTTTGGAGATAAGATCCTGTTGAATTTCTTCTGTTTAAAGGCTTCTCAGGATTTT 15433
1009 TCAGGCAAGAGGGGTTCAGGCCAGGCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAG 1068
15432 TCAGTGGAAAATTTTCTAAGGTTTATAGTTTGGGTGAAAATCTGGTTTACGAATCGTG-G 15374
1069 GAGGTACCTGACCCGGCATGTTTGGTGAGGGAGATTCAGGAAGTCTTCTCGAAGAGAGG 1128
15373 TAGCTGAAGCCAGAACAAAGATAATCTTGATCAAGGGTATGTAACTCATGAGGAAAG 15314
1129 CTGTCGGAGCTGAGACTCATAGATGAGTGGGAGGAGGTTTCC---AGCAGAAAGACCA 1185
15313 AAAAAAGAGCTTAAATGAGAAAGGTTGATATAACAGAGAGACAAGATGGAAGAGCATT 15254
1186 GCACCTACAAAAGCATGACTTTGAGAGAGCATTCATCCATTCAACTGATGAATTTTCAG 1245
15253 TTTTTTTTTTTTTTTTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15194
1246 ACTGGGCAGCTGGCTCATGCTGTAAATCCAGCACTTTGGAAGGCTGAATGGGAGGAT 1305
15193 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15134
1306 GACTTGAGCTAGGCATTTGTGACAAAGCCTGGGCAACATGTTGAGAGACCTGCCTCCACA 1365
15133 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15074
1366 AACAACAAACAAACAAACAAATATATACCTGGTACCATGGTACCAGGTACATAGAAA 1425
15073 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 15014
1426 TCACCTCAGCAGATATGTTGTCTCTCTACTGTGGGAGAGCGGCTTTATATCTGAGTA 1485
15013 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14954
1486 AGACAATAGAGGAGGGAATATAATCTTAAATAGAGAGTAAGATTTGAGAGCAACAC 1545
14953 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14894
1546 AGGCACAGGCATATCTACAGGGTAAAGAGGAAATCAGGGAAGGCTTCTCAGAGAGGT 1605
14893 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14834
1606 GACATTTAAGCCGGGACATGAAGGATGAACAGATTAGTTTCAACAAGGATGGGATGAAA 1665
14833 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14774
1666 GGGTGAGAGTGTGAGGAGGAGGAACTGAGAGTATAGGCTTACAGCGGATCCTG 1725
14773 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14714
1726 AGCCCTTCAGAGAGTGTGAGAGGAGACACAGCGCAGTCTGTAGTGGGTTAAGTAAACA 1785
14713 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14654
1786 AGAAGCCAGGAAATCCCTGGTCAATGAGGCGCTGTGAGTCAAGTGTGTTGGGCT 1845
14653 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14594
1846 TTTGTTTTTCTGGGAGAGTGCATTTTAAAGCGGAAACAGCTGTATTTCAGAGTTGGGAA 1905
14593 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14534
1906 GATCTGTGTTGCTGCTGAAAGGAGTGAACCTGGAGGCTAGGAGCCAGGCTGATAGG 1965
14533 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14474
1966 GAGGATCCAGGAGTGTGAGGAGGCTGGGAGGCTCCGCGGTGATGGAACAGGCTGGGCGCA 2025
```


RESULT 8

US-09-880-107-1542/c
; Sequence 1542, Application US/09880107
; Patent No.. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Dargi T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1542
; LENGTH: 110096
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AD000092
US-09-880-107-1542

Query Match 8.3%; Score 247.6; DB 9; Length 110096;
Best Local Similarity 47.3%; Pred. No. 2.3e-63; Indels 57; Gaps 11;
Matches 1149; Conservative 0; Mismatches 1224;
QY 287 CAATGGAGTGTAAACAGTGCCTCAGCGGGTGCAGTGGCTGAAATCCCAA 346
DB 58480 CAAACAAAAGAAACAGTGAAGCTGCACTGGCGCAGGTGGCTCAGCGCTGTAATCCCGAG 58421
QY 347 CACTTTGGAGGGAGGTGGTGGATCACTTGAAGTCAAGAGTTCAGACCAAGCTGGC 406
DB 58420 CAAATTTGGAGGCCAAGGAGGTGGATCACTAAGTCAAGAGTTCAGACCAAGCTGGC 58361
QY 407 CAACATGGTGAACCTGCTCTACTAAATAATATAAATAATAGCTGGCATGGTGGG 466
DB 58360 CAACATGGTGAACCTGCTCTACTAAATAATATAAATAATAGCTGGCATGGTGGG 58301
QY 467 TACCTGTATCCAGATACCTTGGAGGTGGAGGAGAAATCGCTTGAACCCGGAGGC 526
DB 58300 TGCCTGTATCCAGCTACTCAGAGGCTGAGATGGGAAATCACTTGAACCTGGGAGAT 58241
QY 527 AGATGTTCAGTGAACCAAGACTGTCCTCACTGCTCACTGCTGGGCAACAGAGTGAGCC 586
DB 58240 AGTGGTTCAGTGAGCGAGATCGTGGCACTGCTCACTTCCATCCTGGCAGCTGAGCGAGAC 58181
QY 587 TCCATCTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 646
DB 58180 CTCGTCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 58121
QY 647 AGGACTGAGCCTTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 706
DB 58120 GCACCTGTATCCAGCTACAGGGAGGTGAGTGGGAGAAATGCTTGAACCCAGAGG 58061
QY 707 ACCATCAATGATCCAGTGAAGCTGAGGTTCAGAAAGATGCGGAGCGCTTCAAGG 766
DB 58060 TGGAGGCTCAGTGAGCCAGATCATGCCACTGCACTCCAGCTGGGCGACAGGTGAGA 58001
QY 767 TCGTGGGATGTTGGTGGGCAAGCCCTCGAATAATAGAAACAGTTCCTGTATTACAAACA 826
DB 58000 CCCCTTCTCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 57941
QY 827 AAGCAGGAGCCCATGCTGGGTGCTGCGAGGAAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 886
DB 57940 GGG-TGGAGGCCAGGGCCCTGGGAAGCCCTTGTATGCTGTGTGAAGGAATTTAGACTTGAT 57882

QY 887 GCTGCTTCCCGAGCGCACCCTAGGCGAGTGGGAAACAG-----ACTCACCACACAGT 938
DB 57881 CTTGTGGCTGGTGGGATGGTTTTAGGCTAACAAACAGAGTAGGAATCAGATTACAGT 57822
QY 939 CCCAGCCAGAGTGGTCAAGGCAAGATGGGCAAGCAGCGGA---GAAAGGTCAAGGTTG 995
DB 57821 GTAGAAATGCTACTAGGGCTGACACTGTGGCTCAGCTTGTATCTCAGCACTTTGGGG 57762
QY 996 GGATGGGAGGGTTCAGGCAAGAGGGTTCAGGGCTAGGGCTAGGGAAGCCCTGGGACTG 1055
DB 57761 CCGAGGCAAGATTTTTGAGGCCAGGAGTTTGAAGCAGACTTGGGTAACATAGCAAGACC 57702
QY 1056 TAGGAAATTTAGAGAGGTACCTGACCCCGCATGTTTGGTGAAGGAGATTCAGAAAGTCTT 1115
DB 57701 TCCATCTATACAGAAAAATTAAACAAATTTAGCCAGAGTAGTGGTACCCCTGTAGT 57642
QY 1116 CCTGAAGAGAGGCTGTGGAGCTGAGA--CTCATAAGATGATGGGAGGGTGTTCAG 1173
DB 57641 CCTAGATATTCAAGAGGCGAGAGCAGAGGATCACTTGAGCCTGGGGAGTTCAAAGTTG 57582
QY 1174 GCAGAAAGCAGCAGCCTTACAAAAGCATGCTTTGAGAGAGCATTCATCATTCACTG 1233
DB 57581 CAGTAAATGTACTGATGTACTCCAGCTTGGCAATGTAGCAAAACCTCGCTTCAAAAGAAA 57522
QY 1234 ATGAATTTTCAGACTGGGCAAGCTGCTCATGCTGTAATCCAGCACTTTTGGAGGGCTG 1293
DB 57521 AAAAAGGTCAAGCCGGTGGCTCAGCCCTGTAATCCAGCACTTTTGGAGGGCG 57462
QY 1294 AATGGGAGGATGACTTGAAGCTTAGGCTTTGTGACAAAGCTTGGGCAACATGGTGAGACC 1353
DB 57461 AGAGCGGCGAT--CATGAGTCAAGATGAGAGCATTCAGCACTTCCTGGCTAAACATGGTGAAC 57404
QY 1354 CTGCTCTCAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1413
DB 57403 CGCCTCTACTAAAAAATACAAAAATTTAGCCAGGTGTGGTGGCGGCGCTGTAGTCCCA 57344
QY 1414 GGTACATAGAAATGACTCAGGCAGATATGGTCT---CCTCTCTACTGTGGGAGAGCGG 1470
DB 57343 GCTACTCGGAGGCTGAGCAGGAGAAATGCAATGAACCGGAGGAGAGCTTGCAGTGA 57284
QY 1471 GCTTATCTGCACTAAGCAATAGAGGGAGGAAATATATCTCTAAATAGAGGTACAGA 1530
DB 57283 GCGAGATTTGCTCCACTGCACTCCAGCTCGGGGAGCAGAGCAGACTCCATCTCAAAAAA 57224
QY 1531 TTTGAGAGCAACACAGGCGCAGGCATATGTACAGGGTAAAGAGGGAATCAGGGAAG 1590
DB 57223 AATATATATATAAAAGGTCAAGCGGTGGTGGCTCATGCTATAATCCAGCACTTTGG 57164
QY 1591 CTTCTCAGAGAGGTGACATTTAAGCGGACATGAAGGATGAACGAGTTAGTTCACCAA 1650
DB 57163 AGGCGAAGCAGCGGA---TCACTTGAGTCAAGAGTTTGAGACCGAGCTGCGCAACAT 57107
QY 1651 GGATGGGATGGAAGGGGTGAGATGATGAGGCGAGAGGAAATCGCAGGATCATAGCCCT 1710
DB 57106 GGTGAGACCTCTCTACTATAAAATATAAAATTTAGCCGCGCTGGTGGTGTGCTGCTG 57047
QY 1711 AGACAGGGATCCTCAGCCCTTGAGGAAGTGAAGAGAACACAGCGCACTCGTAGTGGGT 1770
DB 57046 TAGTCCCACTACTCAGGAGGCTGGGCGACAGAAATGACTCAGCGCCAGAGGTGAGGT 56987
QY 1771 TAAGTAAACAAAGCTGAGAGCCAGGGAATCCCTGTCTATGAGGGCCTGTGAGTCACGT 1830
DB 56986 GSCAGTGAGCTGAGATTGCAACCACTCCAGCTGAGCGACACTCCAGCTGTCACTCCAGC 56927
QY 1831 CAGAGTGTGGGCTTTGTTTTTCTGGGAGCAGTCTGATTTTAAAGCAGGGAACAGCTGT 1890
DB 56926 CTGAGCGACCGAGTGAAGTCTCTCAAAAAACAAATATAATACAAA-----A 56876
QY 1891 ATTCAAGTGTGGAAGATCCTGTGGTGTGCTGCTGAAGGGGATGAAACTGGAGGCTAGGA 1950
DB 56875 TTTTTCACAAAGAAATTTCTTCAAAAGAAATAAATTTTAAACATAAATTTATT 56816
QY 1951 GCCCAGGGTGATAGGAGGATCCAGGGTGTATGGGAGGCTGGAGGTCCCGGTGATGGA 2010

Db 56815 ACAAAATTTTTTACAAAATAAATAAGTAATAATAAAGTCCCTATGSCCTAGTTGA 56756
QY 2011 CCAGGGCTGGGCCAGGGATGGGAGAGAGTAATTTGGAGAGGCGCTGGGCTCTGG 2070
Db 56755 GAAGTGAAGATTATAGAGCTCTAATGATGTGTAGATTTAGAGACATTAAGAAGGTAGGGC 56696
QY 2071 CCGAGGAATGATGGTGGCTGAAACAGGAGAGAGAGATGCTTAGCCACTTTGGNAC 2130
Db 56695 CGGCAATGGTGGCTACGCCCTGTATCCAGACCTTTGGGAGGCTTGAGGCAAGGTAGA --- 56639
QY 2131 ACAGTAGGCAAGGACAGGACACCCCAAGGGGAAGTCCCAAGAGACACGACAGGCTG 2190
Db 56638 TCAGAGGTGAGGAGTGGAGACAGCCTGCGCAAGATGTTGAACCCCGTCTCTACTAA 56579
QY 2191 GCATTGACAGGGAAGTCTCTCTGAGCAGGTGTCTTGGATAGAGGAGGAATAATGTGC 2250
Db 56578 AAATACAAAAAATTAGCTGGGTGTGGTGGCAGTGCCTGTAATCCAGCTACTTTGGGAGGC 56519
QY 2251 AGTTCCATCCCTCCCTCTCTCTGTTCAACCTTAACTACATGGGACAGACCCAGTG 2310
Db 56518 CGAGGCAAGAGATCGC-----TTGAAACCGGAGGTGGAGGTTCAGTACCTGAT 56465
QY 2311 GGACTCCATAAATGATGGGTGGTGGATGGAAGGAAGGAGGAAACAACCTTTTCAT 2370
Db 56464 CGCTCCACTGCCTCTAGCTTGGG-----CAACAAGCAAGACTCTGT 56422
QY 2371 TCATCTGGTTATTTACAGACAGGCGCAGGTGCGGTGCTCAGCTTGCCTATTTAGCACT 2430
Db 56421 CTCAAAAACAATAAATAAAGGCTGGGCACAGTGGTTCATGCTGTAATCCAGCACT 56362
QY 2431 TTGGAGGCTGAGTGGGTGATTTACCTCAGGTGAGGATTCAGACAGGCTAGACAAC 2490
Db 56361 TTGAGAGCCGAGGAGGCGGATCACCCTGAGGTGGGAGTTTGAGACCGCCTGACCAAC 56302
QY 2491 GTAGAGAAACCCATCTCTACTGAAGATATAAATAAGCTGGGCTAGTGCCATATGCT 2550
Db 56301 ATGGAGAACTCCGTCTCTACTAAATAACAAATTCGCCAGGTGTGGTATGCTCT 56242
QY 2551 GTAATCCAGCTAGTCGGGAAGCTGAGGAGGAGAAATCGTTGAACCCGAGAGCAGAGG 2610
Db 56241 GTAATCCAGCTATTCGGGAGGCTGAGGCAAGAGAAATCGCTTGAACCCGAGGCGGAGG 56182
QY 2611 TTGCGGTGAGCTGAGATCGTCCATTCCTCAGCTGAGGTCGAGCAAGCAAGCTCGT 2670
Db 56181 TTGAGTGAAGCAGATCGCGCATTCGATTCAGCTGAGGTCGAGGTCGAGCAAGCAAGCT 56122
QY 2671 CTCATAATAATAATAATTACAAACAGAA 2700
Db 56121 TCTGAAAAAATAAATAAATAAATAA 56092

RESULT 9
US-10-412-277-3/c
; Sequence 3, Application US/10412277
; Publication No. US2003017591A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001067DIV
; CURRENT APPLICATION NUMBER: US/10/412,277
; PRIORITY FILING DATE: 2003-04-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(786431)

OTHER INFORMATION: n = A, T, C or G
US-10-412-277-3
Query Match 8.2%; Score 245.8; DB 15; Length 786431;
Best Local Similarity 84.2%; Pred. No. 2.3e-62;
Matches 277; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 282 TCTCTCAATGAGATGATTAACAGTGCCTCATCAGCGGGTGCAGTGGCTAGTCCCTGAAT 341
Db 572672 TTTTTCATTACTCTATTAAACAGTGACTAGAGGCGCAGGTGCTGTGGCTCATGCTATAAT 572613
QY 342 CCCAACACTTTGGAGCGGAGGTGGTGGATCACTTTAGGTGAGGATTCGAGACCGAGC 401
Db 572612 CCCAGCACTTTGGAGCGGAGGTGGTGGATCACTTTAGGTGAGGATTCGAGACCGAGC 572553
QY 402 CTGCGCAACATGTTGAAACCCCTGCTCTACTAAAAATATAAAAAATTAGCTGGCATGGT 461
Db 572552 CTGCGCAACATGTTGAAACCCCTGCTCTACTAAAAATATAAAAAATTAGCTGGCATGGT 572493
QY 462 GTGGTACTCTGTAATCCAGATCTTTGGAGGTTGAGGAGGAGAAATCGCTTTGAACCGG 521
Db 572492 GCGTACACCTGTGATCCAGCTACTTTGGAGGTTGAGGAGGAGAAATCACTTTGAACCTGG 572433
QY 522 GAGGACATGTTGAGTGAACCAAGACTGTGCCACTGCCTCAGCTTGGCAGACAGAT 581
Db 572432 GAGTGGAGGTTGAGTGAACCAAGACTGTGCCACTGCCTCAGCTTGGGCTGGTGCAGAT 572373
QY 582 GAGCTCCATCTCAAAACAAACAAACAAA 610
Db 572372 GAACTCCGCTCAAAACAAACAAACAAA 572344

RESULT 10
US-09-997-722-154/c
; Sequence 154, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; PRIORITY FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIORITY FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 154
; LENGTH: 96594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-997-722-154

Query Match 8.2%; Score 245.6; DB 12; Length 96594;
Best Local Similarity 48.2%; Pred. No. 8.7e-63;
Matches 1158; Conservative 0; Mismatches 1174; Indels 68; Gaps 14;
QY 314 GCGGGTGCAGTGGCTAGTGCCTGAATCCCACTTTGGGAGGCGGAGGTGGTGGAT 373
Db 17837 GCGGGGCTGGTGGCTTATGCTGTAAATCCAGCACTTTGGGAGGCGGAGCGGAGAT 17778
QY 374 CACTTGAAGTTCAGGAGTTCGAGACCACTGGCCAACTGGTGAACCCCTGCTCTACTA 433
Db 17777 CACTTGAAGTTCAGGAGTTCGAGACCACTGGCCAACTGGTGAACCCCTGCTCTACTA 17718
QY 434 AAAATATAAATAATAGCTGGCATGGTGGTGGTGAATCCAGATCTTGAATCCAGATCTTGGAGG 493
Db 17717 AAAATATAAATAATAGCTGGCATGGTGGTGGTGAATCCAGATCTTGAATCCAGATCTTGGAGG 17658
QY 494 TTGAGCAGGAGAAATCGCTTGAACCCGAGGAGGAGATGTTGAGTGAACCAAGCTGTGC 553

Db 17657 CTGAGCATGAGAAATCTCTTGAACCCAGCTGGTAGAGGTTGCAGTGAGCCCAAGATCATAC 17598
Qy 554 CACTGCACCTCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAAAACAAACAAAAAGC 613
Db 17597 CACTGCACCTCCAGCTGGGCAACAGAGTGAGCTCCATCTCAAAACAAACAAAAAGC 17541
Qy 614 AGTGCCCATCATGTAGGATTGAGTGATTTGAGTGAGGACTGAGCCTTTGCAAGTGAGCA 673
Db 17540 -----CTTCTCTGGGTAGGCAAAAGGGCTGGGGAGTGAATTAATCACCAATGGCCAA 17487
Qy 674 CTCCTAATC-ACCAGGTTGTAGTATCAGTGATAACCATCAATGATCAGGTAAAGCCCT 732
Db 17486 TTAATTAATCAATCATGTCTAGTGTAGGAGGCTTACCTAAACCCAGNAGATAGGCT 17427
Qy 733 GAGGTTTCAGAAAGATGCGGAGCCCTTTCAAGGTGCTGGGATTTGGTGGCAAGCCCTC 792
Db 17426 CAGGAGCTTCTAGGTTGGTGAGAACATGTAGAGATGGTGGAGGCTGGTGTGCTGGAG 17367
Qy 793 GNAATATAGAACAGTTCTCTGTATTAAACAGAAAGCAGGAGGCCCATGCTGGGTGCTG 852
Db 17366 AGGCGATGGAAGTGCACACCCCTTCCCATACATACCTTGTCTGTGTATTTCTTCAGCTG 17307
Qy 853 CCAGGAATCTCAGTAGTAACCTAAGACAGCACCGGTGCTGCTCCCGAGCCACCTAGGCCA 912
Db 17306 GCTGCTTCTGAGTTGTTTATATATAGCTGATTAATCTAGTAAGGAATGTTCTGTGACTG 17247
Qy 913 GTGGGAAACAGACTCACACAGTCCAGGCCAGAGTGTCTCAGGCCAGGCAAGATGGGAA 972
Db 17246 AATGTGGAGGACTTACCACTTGTGCTTGTAGTCTGAAGTGGGGGTCTGGAATGGGGTG 17187
Qy 973 GCACGGGAGAAAGTTCAGGTTGGATGGGAGGGTTCAGGCAAGAGGGGTTCAGGCCA 1032
Db 17186 GACGGGAGGACTTAACATTTGGGCTCTGCACTAACCTGGCAGTGTAGTGTACAGCTG 17127
Qy 1033 GCCTGAGGGAAGCCCTGGGACTGTAGGAATTTAGAGGAGTACCTGACCCCGCATGTTTG 1092
Db 17126 AGCTAAATGTAGGACACCCAGTGTCTGCTAAAGACTGAATTTGTGGGAAA----- 17073
Qy 1093 GTGAGGAGATTCAGGAAGTCTTCTGGAGAGAGGCTCTCGAGCTGAGACTCATAGA 1152
Db 17072 ---AAAAACACCACACATTTGGTGACAGAGTGAAGGAATTAATATATGGGTACAGAGTAG 17016
Qy 1153 TGAGTGGGAGGAGTGTTCAGGCAGAAAGACCCAGCACCTACAAAAGCATGACTTTGAGAG 1212
Db 17015 AGAGGAATGAGGTTTCTTTACAGTGGTATAGCA-----TAGAATG 16973
Qy 1213 AAGCATTTCAATCCATTCATGATGAATTTTCAGACTGGGACGCTGGCTCATGCCGTAA 1272
Db 16972 AAATATTCAGCCCTTTAAAGAAAAAATTT--GACTAGATGAAGTGGCTCATGCTGTAA 16915
Qy 1273 TCCAGCACTTTGGAAAGGCTGAATGGGAGGATGACTTTCAGCTTAGGCAATTTGTGCAAG 1332
Db 16914 TCCAGCACTTTGGGAGGCCAAGGCAGGAGTATCGCTTGGCCCGAGGTTTGAGACCAG 16855
Qy 1333 CTTGGCAACATGGTTCAGACCC--TGCCCTCCAAAAACAAACAAACAAACAAATCAT 1390
Db 16854 CTTAGGCAACAAAGCAAGCCCGATGATCTACAAAAATTAAGGTTAGCCAGCATGCT 16795
Qy 1391 TATACCTGGTACCATGGGTACCAAGGTACATAGAAATGACTCAGGCGAGATATGGTGTCTC 1450
Db 16794 GGTACACTCTGTGTGTCTCACCTTCTCGGAGGCTGAGGCAAGGATCATCTGAGCCCA 16735
Qy 1451 TCTACTGTGGAGAGCGGGCTTACTGTGAGTAGAACAATAGAGGGGGAATATAAT 1510
Db 16734 GGAGTTTAGAGGCTGCGAGTGAACCAACATCTTCTGCACTCCAGCCCTGGATAGCTGAG 16675
Qy 1511 CCTAAATGAGAGGTACAGATTTGAGAGCAACACAGGGCACAGGCAATGTACGAGGCT 1570
Db 16674 CAAGACCTCTCTGAAAAAATAATAAATAACAATTAAGTTAAATAAATAAATA 16615
Qy 1571 AAAGAGGGAATCAGGGAGGCTTCTCAGAGAGGCTGACATTTAAGC-----GGGACATG 1525
Db 16614 AAAAAAATCTCTGCAACATGAGATGACAGATGAACCTTGAAGCCATTTAAGCCAAAGTG 16555

Qy 1626 AAGGATGAACAGATTAGTTTCAACAGGATGGATGGAAGGGGTGAGAGTGATGAGGCA 1685
Db 16554 AAGTAAGCCATCTCAGAGCAAAATTAATACATGATTCCTCAACAGAGGTATCTAAATA 16495
Qy 1686 GAGGAACTGAGGATCTATAGGCCCTAGACAGGGGATCTCGACGCCCTTTGAGGAAGTGAGA 1745
Db 16494 GGCAACTCATAGAAGCAAGAGGGGCACAGTGGTTGCCAGGGTAGGGCTTTTCAGGAAA 16435
Qy 1746 GAAGACAGCGGAGTCTGTAGTGGTTAAGTTAAACAAAGCTGAGAACCCAGGGAATCCCTG 1805
Db 16434 TGGGAGTTGTTAATCAAAAAAGTATAAAGTTTCAGTAATGCAAAATGAACGCAATCTAGA 16375
Qy 1806 GTCATGAGGGCTCTGTAGTCAAGTCAAGTGTGTTGGGCTTTGTTTCTCTGGGAGCAG 1865
Db 16374 GATCGCCATCAACATTTGTGC-TGTGTGTTACTATACATGATAAACTCTCTTAAAGG 16316
Qy 1866 TCGATTTTAAAGCAGGAAACAGCTGTATTACAGATTTGGGAAGATCTGTGTGTTGCTG 1925
Db 16315 TTAAGAAAAACAATTTAAAAAAATTTTTTTTTTAGAGACAAGGCTTGTCTCTCACTCAG 16256
Qy 1926 AAGGGATGAACCTGAGGCTAGGAGCCAGGCTGATAGGAGGATCCAGGGTGTATGGG 1985
Db 16255 GCTAGAGCGCAGTGGCACAATCATAGTTCACTGCAAGCTCAGACTCTTAAGCTCAAGTA 16196
Qy 1986 AGGCTGGGAGGTCGCCGTGTATGGACCAAG--GCTGGGGCCAGGGGATGGGAGGAAG 2042
Db 16195 TCCTCTGCTCAGCTCTGGAGAACTTGGAACTACAGGTGATGACATCATGCTCAACT 16136
Qy 2043 AGTAATTTGGGAGAGGCTGGGCTCTGGCCGAGGAATG-----GATGTGGGCTGAAC 2096
Db 16135 AATTTTTTAAATTTTTTATAGAGACAAGTCTCCATTTGTGCAGACTGCTTGAATCTC 16076
Qy 2097 AGGAGAGAGAGATGCTTTAGCCCACTTTGGAACACAGTACAGGAGGAGGACAGACACC 2156
Db 16075 CTGGCTCAAGAGATCTTCTCTCAGCTCCCAAGTGTGGGATTCAGAAAGTGAGC 16016
Qy 2157 CAAGGGAGTCCCAAGAGACCAAGCAGGCTGGCATTTGGACAGGGAAGGCTGTGTG 2216
Db 16015 TACCATGCCAGCCAAATTAATCTATTAAGAAAAAAGAAAGAAACATATATCT 15956
Qy 2217 AGCAGGTGTCTTGGATAAGGAGGAAATGTTGTC-----AGTTCCATCTCTCTCC 2267
Db 15955 ACAACGTGGAAGAACTCTCAAAACATATAAGCGCAACTAGAAAGACTCAGTAGAATGCA 15896
Qy 2268 TCTCTGTTCAACCTCTAACTACATGGGCAAGGACCCAGTGGGACTCCATAAATGATG 2327
Db 15895 TATAGTATTATCCCATTTACAGAGCTCAAAACAGGCAAACTGAGCTATATTATTAT- 15837
Qy 2328 GGATGGGTGGATGGAAGGAAGGAGGAAACAACTCTTTCATTTCACTCTGTTATTAT 2387
Db 15836 AGATGTATAGTACAGGCAACTATGAACACAAGAAAGCAATTAACAAAAATTTGGAC 15777
Qy 2388 AGAACAGGCCAGGTGGGTGCTCAGCTTTGCCATTTAGACCTTTGGGAGGCTGAGGTG 2447
Db 15776 AATAGGCCAGGAGCGGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGCCAGGCGAG 15717
Qy 2448 GTGGATTACCTCAGGTTCAGAGTTCAAGACCAAGGCTAGAACAGTACAGAAACCCCATCT 2507
Db 15716 CGAGATTACCTGAGGTTCAGAGTTTAAGACCAAGCTGGCCCAATGATGTAACCCCGTCT 15657
Qy 2508 CTACTGAAGAT--ATAAAATTAGCTGGGCTAGTGGCATATGCTCTTAATCCAGCTAGT 2565
Db 15656 CTACTAAANTAGAAAAATTAGCCAGCCATGTGTGCGAGCGCTGTAATCCAAGTACT 15597
Qy 2566 CGGGAAGCTGAGCAGGAGAAATCGCTTTGAACCCGAGAGCAGAGGTTGCGGTGAGCTGAG 2625
Db 15596 TGGGAGGCTGAGCAGGAGAAATGCTTTGAACCCAGGAAGCAGAGGTTGAGTGAGCGGAG 15537
Qy 2626 ATGTCGCCATTCACCTCCAGCTGGGTGACAA-AGCAAGACCTCTGCTCTCAATTAATAA 2684
Db 15536 ATCGGCCACTGCACTCCAGATGGCAACAGAGCAAAACTCCGCTCTCAAAATAATAA 15477

; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 256849
 ; LENGTH: 546
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-256849

Query Match 8.1%; Score 243; DB 13; Length 546;
 Best Local Similarity 84.0%; Pred. No. 3.6e-63;
 Matches 273; Conservative 1; Mismatches 51; Indels 0; Gaps 0

Qy	287	CAAAATGGAGTGATACAGCTGCCCATCAGCCGGGTGCAGTGGCTAGTGCCTGAAATCCCAA	346
Db	541	CACAGGGTTTAAACAAGCATATAGGCCCCGGAGCCATGGCTACGGCTGTATATCCCAG	482
Qy	347	CACTTTGGGAGCGGAGGTGGTGGATCACATTGAGGTCAGAGTTCGAGACCAGCCTGGC	406
Db	481	CACTTTGGGAGGCTGAGTGGTGGATCACATTGAGTTCAGAGTTCGAGACCAGCCTAGC	422
Qy	407	CAACATGGTGAACCCCTGTCTCTACTAAAAATATAAAAAATTAGCTGGGCATGGTGGTGGC	466
Db	421	CAACATGGCRAAACCCCATCTCTACTAAAAATACAAAAAATTAGCTGGGCATTGTGGCGCA	362
Qy	467	TACCTGTAATCCCAAGATACACTTGGGAGGTGAGGACGAGAGATCCGTTGAAACCCGGAGGC	526
Db	361	TGCTGTGAATCCCAAGTATTTTGGGAGGCTGAGGACGAGAGATCCGTTGAAACCCAGGAGGC	302
Qy	527	AGATGTTTGCAGTGAACCAAGACTGTGCCACTTGCACTCCAGTCTGGGCAACAGAGTGAGCC	586
Db	301	AGAGTTTGCAGTGAGCCAAAGATTACGCCACTGCACTCCAGCTGGGCAACAGAGCAAGAC	242
Qy	587	TCATCTCAACCAACAAACAAAAA	611
Db	241	TCCATCTCAACCAACAAACAAAAA	217

RESULT 13
 US-10-027-632-256849/c
 ; Sequence 256849, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 256849
 ; LENGTH: 546
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-256849

Db 884 GCAGAGATCGAGATCGTGCACATGCACTCCAGCCTGGGCAACAGAGTGAGACTCTGTCT 943
Qy 594 CAAACAAACAAACAA 611
Db 944 CAAAAAAAAAAAAAAAAA 961

Search completed: September 18, 2004, 01:40:19
Job time : 1348.64 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 07:12:45 ; Search time 7312.95 Seconds
(without alignments)
12254.481 Million cell updates/sec

Title: US-10-669-693-3_COPY_1980_4980

Perfect score: 3001

Sequence: 1 cttcttcctccctaccccc.....atcctggtaacacagtga 3001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_esti:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241.8	8.1	757	28	AQ528478
C 2	241.6	8.1	760	28	AL691744
C 3	240.2	8.0	478	28	B67141
4	240	8.0	808	28	BZ611126

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C	5	239	8.0	833	28	BZ599987
	6	237.6	7.9	553	14	CA438009
	7	237.4	7.9	551	12	BM754642
C	8	237.4	7.9	657	14	CF123536
C	9	237.2	7.9	696	13	BU616112
	10	237	7.9	690	28	AQ415537
	11	236.8	7.9	494	10	AW471332
C	12	236.8	7.9	932	13	BQ860197
	13	236	7.9	998	28	AQ45514
	14	236	7.9	656	29	AG035209
	15	235.8	7.9	466	28	AQ590440
C	16	235.6	7.9	470	28	AQ226326
	17	235.6	7.9	588	13	BA488003
	18	235.6	7.9	677	29	AG169901
	19	235.4	7.8	412	10	BE062478
	20	235.2	7.8	490	10	BE281645
	21	235.2	7.8	746	12	BG774140
C	22	235	7.8	560	10	BE155951
C	23	235	7.8	724	14	CB555645
C	24	235	7.8	995	12	BM423099
	25	234.8	7.8	477	28	AQ221138
	26	234.8	7.8	611	14	CD704731
C	27	234.6	7.8	501	13	BA485916
	28	234.4	7.8	672	13	BA501311
	29	234.2	7.8	675	13	BU664420
C	30	234.2	7.8	3010	11	BC032827
	31	234	7.8	416	10	BE062476
C	32	234	7.8	614	28	AQ627870
C	33	234	7.8	678	28	AZ518835
C	34	234	7.8	687	14	CA448731
	35	233.6	7.8	588	28	AQ45949
	36	233.6	7.8	877	28	AQ739838
C	37	233.2	7.8	521	13	BA485089
	38	233.2	7.8	642	28	B59854
	39	233.2	7.8	648	28	AQ266668
	40	233.2	7.8	648	28	BZ611349
	41	233.2	7.8	685	28	AQ423897
	42	232.8	7.8	439	9	AL712324
C	43	232.4	7.7	439	9	AA515728
C	44	232.4	7.7	463	28	AQ535344
	45	232.4	7.7	541	28	BZ604634

ALIGNMENTS

RESULT 1	AQ528478	757 bp	DNA	linear	GSS 18-MAY-1999
LOCUS	RPCI-11-349N11-TV	RPCI-11	Homo sapiens	genomic clone	
DEFINITION	RPCI-11-349N11, genomic survey sequence.				
ACCESSION	AQ528478				
VERSION	AQ528478.1	GI:4840591			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 757)				
AUTHORS	Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.				
TITLE	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building				
JOURNAL	Unpublished (1997)				
COMMENT	Other GSSs: RPCI-11-349N11.TJ Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbeetigr.org Clones are derived from the human BAC library RPCI-11. For BAC				

library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.

FEATURES

Location/Qualifiers
1..757
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7633954"
/db_xref="taxon:9606"
/clone="RPC1-11-349N11"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPC1-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPC111 Human Male BAC Library"

ORIGIN

```
Query Match      8.1%; Score 241.8; DB 28; Length 757;
Best Local Similarity 87.7%; Pred. No. 2.7e-43;
Matches 264; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 314 GCCGGTCAGTGGCTAGTGCCTGAATCCCAACACTTTGGAGCGGAGGTGGTGGAT 373
DB 268 GCCTGGTCAGTGGCTCATGCTGTAATCTCAGCACACTTTGGAGGCTGAGTGGTGGAT 327
QY 374 CACTTGAGTCAAGGATTCGAGACAGCAGCTGGCCAAACATGGTGAACCTGTCTCTACTA 433
DB 328 CACTTGAGTCAAGGATTCAGACAGCAGCTGGCCAAACATGGTGAACCTGTCTCTACTA 387
QY 434 AAAATATATAAATTAGCTGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 493
DB 388 AAAATATATAAATTAGCTGGGATGGTGGGAGGACCTGTGAATCCCACTTGGGAGG 447
QY 494 TTGAGGAGGAGATCGCTTGAACCCGGAGGACATGTTCAGTGAACCAACAGCTGTGC 553
DB 448 CCGAGGAGGAGATTTGCTTGAACCCGGAGGATGTTCAGTGAACCAACAGCTGTGC 507
QY 554 CACTGCCTCCAGTCTGGGCAACAGAGTGAGCTCCATCTCAACAAACAAACAAAGC 613
DB 508 CACTGCCTCCAGTCTGGGTACAGAGTGAGACTCTGCTCAAAAAAAGGAGAGAGAGA 567
QY 614 A 614
DB 568 A 568
```

RESULT 2

AL691744/c
LOCUS B67141 760 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp313M0830.s1.313 (synonym: hlcc2) Homo sapiens cDNA clone genomic survey sequence.
ACCESSION AL691744
VERSION AL691744.1 GI:19617321
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 760)
AUTHORS Koehrer, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.
TITLE EST (Koehrer, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,

Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.

ri sequence also available.
This clone (DKFZp313M0830) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers
1..760
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp313M0830"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="313 (synonym: hlcc2)"
/note="Vector: pRiplex2; Site_1: SfiI; Site_2: SfiI; cDNA-collection"

ORIGIN

```
Query Match      8.1%; Score 241.6; DB 9; Length 760;
Best Local Similarity 85.9%; Pred. No. 3e-43; Indels 0; Gaps 0;
Matches 268; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 300 AACAGTGGCCATCAGCCGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGAGGC 359
DB 323 AAAAGTCACCATAGGCTGGTGGTTCACGCCCGTAATCCAGCACATTTGGAGGC 264
QY 360 GGAGTGGTGGATCATCTTGAAGTTCAGAGTTCAGACAGCAGCTGGCCAAACATGTGAAA 419
DB 263 TGAGGAGTGGATCACCTGAGTCAAGAGTTCGAGACAGCAGCTGGCCAAACATGTGAAA 204
QY 420 CCCTGTCTCTACTAAAAATATAAAATTAGCTGGCATGGTGGTGGTGGTGGTGGTGGTGG 479
DB 203 CCCGTCTCTACTAGAAATACAAATAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 144
QY 480 AGATACCTTGGAGGTTGAGGAGGAGATCGCTTGAACCCGGAGGAGATGTTCAGTGG 539
DB 143 AGTACTTGGAGGCTGAGACAGGAGATCTCTTGAACCCAGGAGAGATGTTCAGTGG 84
QY 540 AACCAAGACTGTGCCACTGCAGTCTGGGCAACAGAGTGAGCTCCATCTCAACA 599
DB 83 AGCCAAGATTTGGCACTGCATCCAGCTGAGCGAGAGTGAGAGTCCATCTCAAAA 24
QY 600 AACCAACAAAAA 611
DB 23 AAAAAAAAAA 12
```

RESULT 3

B67141/c
LOCUS B67141 478 bp DNA linear GSS 21-JUN-1998
DEFINITION CIT-HSP-2017022.TRB CIT-HSP Homo sapiens genomic clone 2017022, genomic survey sequence.
ACCESSION B67141
VERSION B67141.1 GI:2641119
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 478)
AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and Venter, J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other GSSs: CIT-HSP-2017022.TFB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200

Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tldb/hungen/bac_end_search.html
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES

source
 1. .478
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7044072"
 /db_xref="taxon:9606"
 /clone="2017022"
 /sex="Male"
 /cell_type="Sperm"
 /clone_lib="ClT-HSP"
 /notes="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
 HindIII"

ORIGIN

Query Match 8.0%; Score 240.2; DB 28; Length 478;
 Best Local Similarity 87.4%; Pred. No. 5.8e-43;
 Matches 263; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 311 TCAGCCGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGCGAGGTGGTG 370
 |||||
 Db 398 TCAGCCAGGTACATTGCTCATGCTGTATCCAGCAGCTTTGGAGGCCAAGTGGGCG 339
 |||||
 QY 371 GATCATTGAGGTGAGAGTTCGAGACCCAGCTGGCCAAATGATGTAACCCCTGTCTCTA 430
 |||||
 Db 338 GATCATTGAGGTGAGAGTTCGAGACCCAGCTGGCCAAATGATGTAACCCCTGTCTCTA 279
 |||||
 QY 431 CTAATAATATAAATTTAGTGGCGATGTTGCGTACCTGTAAATCCAGACTTTGGG 490
 |||||
 Db 278 CTAATAATATAAATTTAGTGGCGATGTTGCGTACCTGTAAATCCAGACTTTGGG 219
 |||||
 QY 491 AGGTGAGGCGAGAGATTCGCTTTGAACCCGGGAGGAGATGTTGCGTGAACCAAGACTG 550
 |||||
 Db 218 AGGCTGAGGCGAGAGATTCGCTTTGAACCCGGGAGGAGATGTTGCGTGAACCAAGATG 159
 |||||
 QY 551 TGCCTGCACTCCAGTCTGGGCAACAGAGTGCCTTCCATCTCAACCAACCAACAAA 610
 |||||
 Db 158 CACCACTGCACTCCAGCTGGGCAACAGAGTGCCTTCCATCTCAACCAACCAACAAA 99
 |||||
 QY 611 A 611
 |||||
 Db 98 A 98

RESULT 4
 BZ611126 808 bp DNA linear GSS 08-JUN-2003
 LOCUS WHAD87TF Human MCF7 breast cancer cell line library (MCF7_1) Homo
 DEFINITION sapiens genomic clone MCF7_1-21P6, genomic survey sequence.

ACCESSION BZ611126
 VERSION BZ611126
 KEYWORDS GSS.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 808)

AUTHORS Volik.S., Zhao.S., Chin.K., Brebner,J.H., Herndon,D.R., Tao.Q.,
 Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
 Gray,J.W. and Collins,C.

TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)

MEDLINE 12788976
 PUBMED 12788976

COMMENT Contact: Volik SV
 Colin Collins, lab
 UCSF Comprehensive Cancer Center

UCSF Box 0808, San Francisco, CA 94143-0808, USA
 Tel: 415 502 7066
 Fax: 415 502 5665
 Email: svolik@cc.ucsf.edu
 This clone is available from Amplicon Express
 http://www.genomex.com
 Class: BAC ends.

FEATURES Location/Qualifiers
 source
 1. .808
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="MCF7_1-21P6"
 /sex="female"
 /clone_lib="Human MCF7 breast cancer cell line library
 (MCF7_1)"
 /note="Vector: pCBAC1; Site_1: HindIII; This library was
 constructed from MCF7 breast_cancer cell line by Amplicon
 Express (http://www.genomex.com) using their standard
 procedure."

ORIGIN

Query Match 8.0%; Score 240; DB 28; Length 808;
 Best Local Similarity 73.6%; Pred. No. 7e-43;
 Matches 306; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 298 ATACAGTGCCTATCAGCCGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAG 357
 |||||
 Db 210 AGAAATCTACTCTGGGCCGAGCAGTGGCTCATGCTGTAAATCTTAGCATTGGGAG 269
 |||||
 QY 358 GCGAGTGGTGGATCATTGAGGTGAGGTTCGAGATTCGAGCAGCTGGCCAACTGTGTA 417
 |||||
 Db 270 GCCAAGCGAGTGTAGATCATTGAGGTGAGGTTCAGAGTTCAGAGCAGCTGGCCAACTGTGTA 329
 |||||
 QY 418 AACCTGTCTCTACTAAAAATATAAAATTTAGCTGGCATGGTGGTGGCTGCTGTAATC 477
 |||||
 Db 330 AACCTGTCTCTACTAAAAATATAAAATTTAGCTGGCATGGTGGTGGCTGCTGTAATC 389
 |||||
 QY 478 CAGATCTTTGGAGGTTCAGGCGAGAGATTCGCTTGAACCCGGGAGGAGATGTTGCGAG 537
 |||||
 Db 390 CCAGTACTTTGGAGGTTCAGGCGAGAGATTCGCTTGAACCCGGGAGGAGGTTGCGAG 449
 |||||
 QY 538 TGAACCAAGACTTGCCTCAGCTCAGTCTGGGCAACAGAGTGCCTTCCATCTCAAA 597
 |||||
 Db 450 TGAGCCGAGATTGCACCACTGTACTCCAGCTGGCAACAGAGTGCCTTCTCTCAAA 509
 |||||
 QY 598 CAAACAAACAAAGCAGTGCCTCATCTAGTATTGAGTATTGAGTGGAGTGCAGCC 657
 |||||
 Db 510 AAAAAAAAAAAG 569
 |||||
 QY 658 TTGTGCAAGTGCAGCTCACTCAATCACCAGGTTGTAGTATCAGTATACCACTCA 713
 |||||
 Db 570 TTTTTCATCTTGCAATCAATCACCAGGTTGTAGTATCAGTATACCACTCA 625
 |||||

RESULT 5
 BZ599987/c

LOCUS WHACU62TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
 DEFINITION sapiens genomic clone MCF7_1-19K3, genomic survey sequence.

ACCESSION BZ599987
 VERSION BZ599987.1
 KEYWORDS GSS.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 833)

AUTHORS Volik.S., Zhao.S., Chin.K., Brebner,J.H., Herndon,D.R., Tao.Q.,
 Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
 Gray,J.W. and Collins,C.

TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)


```

DEFINITION K-EST0032169 S11SNU1 Homo sapiens cDNA clone S11SNU1-11-A07 5',
            mRNA sequence.
ACCESSION BM754642
VERSION   BM754642.1 GI:19084260
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 551)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
           Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
           Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
           Genome Research Center
           Korea Research Institute of Bioscience & Biotechnology
           52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
           Tel: +82-42-860-4470
           Fax: +82-42-860-4409
           Email: yongsung@mail.kribb.re.kr
           Plate: 11 row: A column: 07
           High quality sequence stop: 551.
           Location/Qualifiers
               1..551
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="S11SNU1-11-A07"
                   /sex="M"
                   /tissue_type="Stomach"
                   /cell_type="Lymphoblast-like"
                   /cell_line="SNU-1"
                   /lab_host="Top10F"
                   /clone_lib="S11SNU1"
                   /note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
                   Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
                   bacterial alkaline phosphatase (BAP) and then decapped
                   with tobacco acid pyrophosphatase (TAP). The decapped
                   intact mRNA was ligated with DNA-RNA linker including SfiI
                   site by treatment of T4 RNA ligase and the first strand
                   cDNA was synthesized with Superscript II using SfiI
                   oligo-dT primer. After first strand synthesis RNA was
                   degraded by NaOH treatment and cDNA was amplified by PCR
                   reaction. The PCR products were digested with SfiI and
                   cloned into DraIII- digested pME18S-FL3 vector. The
                   obtained cDNA vectors were used for transformation of
                   competent cells E. coli Top10F by electroporation method.
                   The cDNA libraries constructed by this method are
                   full-length enriched cDNA library."

FEATURES             source
     source
     1..657
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /tissue_type="CNCAP(3)T-225 cell line"
         /lab_host="DH10B (TI phage resistant)"
         /clone_lib="NIH_MGC_216"
         /note="Organ: Prostate; Vector: pT7T3 Pac; Site 1: EcoR I;
         Site 2: Not I; The library was constructed according
         Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
         1996. Denatured RNA was size fractionated on a 1% agarose
         gel. First strand cDNA synthesis was primed with oligo-dT
         primer containing a Not I site. Double strand cDNA was
         size selected according to mRNA size fraction, ligated
         with EcoR I adaptor, digested with Not I and then cloned
         directionally into pT7T3 Pac vector. The library tag
         sequence located between the Not I site and the polyA tail
         is TAGA. Tissue was provided by Tim Ratliff."

ORIGIN
Query Match      7.9%; Score 237.4; DB 12; Length 551;
Best Local Similarity 87.8%; Pred. No. 2.5e-42;
Matches 259; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 314 GCGGGTGCAGTGGCTAGTCCCTGAATCCACACTTTGGAGGCGGAGGTGGTGGAT 373
Db 209 GTCCGGCGGGTGGCTATACCTGTATATCCAGCACTTTGGAGGCTGAGGTGGCGGAT 268

QY 374 CACTTGAGGTTCAGAGTTCGAGACCGAGCTGGCCACATCGTGAACCCCTGTCTACTA 433
Db 269 CATCTGAGGTGGAGTTCGAGACCGAGCTGACCAATGAGAACCCCGTCTCTACTA 328

QY 434 AATAATATAAATTAGTGGGCAATGGTGGTGCCTACTGTAAATCCAGATCTGGGAGG 493
Db 329 AATAATACAAAATTAGCCGGGATGATGGTGGTGCCTGTATATCCAGTACTTGGAGG 388

QY 494 TTGAGGCGAGAGATCGCTTGAAACCCGGGAGGAGAGTGTTCAGTGAACCAAGACTGTGC 553
Db 389 CTGAGGCGAGAGATCGCTTGAAACCCGGGAGGAGGAGTTCAGTGAACCGAGATTGGCG 448

```

```

QY 554 CACTGCACCTCCAGTCTGGGCAACAGAGTGAGCCTCATCTCAAAACAACAA 608
Db 449 CACTGCACCTCCAGCCTGGGCGACAGAGTGAGACTCCATCTCAAAAAA 503

RESULT 8
CF123536/c
LOCUS     CF123536               657 bp      mRNA      linear      EST 09-SEP-2003
DEFINITION UI-HF-CH0.2-aul-b-11-0-UI.r1 NIH_MGC_216 Homo sapiens cDNA 5', mRNA
sequence.
ACCESSION CF123536
VERSION   CF123536.1 GI:33197844
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 657)
AUTHORS   Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
           discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    889548
COMMENT   Contact: Soares, MB
           Coordinated Laboratory for Computational Genomics
           University of Iowa
           375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
           Tel: 319 335 8250
           Fax: 319 335 9565
           Email: bento-soares@uiowa.edu
           Tissue Procurement: Tim Ratliff
           cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
           cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
           DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
           Clone Distribution: Distribution information can be found at
           http://genome.uiowa.edu/distribution/humanfl.html
           The following repetitive elements were found in this cDNA
           sequence: 139-428, >ALU (matched complement) 344-450, >ALU (matched
           complement)
           Seq primer: pYX-5.
           Location/Qualifiers
               1..657
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /tissue_type="CNCAP(3)T-225 cell line"
                   /lab_host="DH10B (TI phage resistant)"
                   /clone_lib="NIH_MGC_216"
                   /note="Organ: Prostate; Vector: pT7T3 Pac; Site 1: EcoR I;
                   Site 2: Not I; The library was constructed according
                   Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                   1996. Denatured RNA was size fractionated on a 1% agarose
                   gel. First strand cDNA synthesis was primed with oligo-dT
                   primer containing a Not I site. Double strand cDNA was
                   size selected according to mRNA size fraction, ligated
                   with EcoR I adaptor, digested with Not I and then cloned
                   directionally into pT7T3 Pac vector. The library tag
                   sequence located between the Not I site and the polyA tail
                   is TAGA. Tissue was provided by Tim Ratliff."

ORIGIN
Query Match      7.9%; Score 237.4; DB 14; Length 657;
Best Local Similarity 86.5%; Pred. No. 2.6e-42;
Matches 262; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 314 GCGGGTGCAGTGGCTAGTCCCTGAATCCCAACTTTGGAGGCGGAGGTGGTGGAT 373
Db 427 GCGGGCGGAGTGGCTACGCTGTAAATCCAGCACTTTGGAGGCGGAGGTGGCGAGAC 368

QY 374 CACTTGAGGTTCAGGAGTTCGAGACCGAGCTGGCCACATCGTGAACCCCTGTCTACTA 433
Db 367 CATCTGAGTTCAGGAGTTCGAGACCGAGCTGGCCACATCGTGAACCCCTGTCTACTA 308

```


ORIGIN									
	Query Match	7.9%;	Score 237;	DB 28;	Length 690;				
	Best Local Similarity	84.2%;	Pred. No. 3.2e-42;						
	Matches 267;	Conservative 0;	Mismatches 50;	Indels 0;	Gaps 0;				
	/cell_type="Lymphocytes"								
	/clone_lib="RPCI-11"								
	/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;								
	RPCI11 Human Male BAC Library"								
Qy	295	GTGATAACAGTGC	CCATCAGCGGGTGC	CAGTGGCTAGTGC	CTGAAATCCCAACACTTTGG	354			
Db	21	GTTTTAAAGAC	ANATACAGCCAGCGGGTGTATGCTGTAAATCCCGACGACTTTGG	80					
Qy	355	GAGCGGAGGTGGT	GGATCACTTGAGGTCAGGAGTTCGAGAC	CAGCGCTGGCCAACTGG	414				
Db	81	GAGCTCAGGAAGT	GGATCATCAGCTGAGGTCAGGATTTCAAGACCAGCGCTGGACAACTGG	140					
Qy	415	TGAAACCCCTGTC	TCTACTATAAAATATATAAAATAGCTGGGCATCGTGGTGGTACCTGTGA	474					
Db	141	CAAAACCCCTGTC	TCTACTATAAAATACAAAAATAGCCGGGAGTGGTGTGCACCTGTA	200					
Qy	475	ATCCGAGATCTT	GGGAGTTGAGGCGAGGAATCGCTTGAAACCCGGAGGACAGATGTTG	534					
Db	201	ATCCGAGCTACT	CGGGAGGCTGAGCGAGGAGAATCGCTTGAAACCTGGGAGCAGAGGTTG	260					
Qy	535	CAGTGAACCAAG	ACTGTGCCACTCCACTCCAGTCTGGGCCAACAGATGAGCCTCCATCTC	594					
Db	261	CAGTGAGCCAA	GATCGTACCACTGCATCTCAGCCTGGGCGACAGATGAACCTCTGTCCTC	320					
Qy	595	AAACAAACAA	CAAAAA 611						
Db	321	AAACAAAAAA	AAAAA 337						

RESULT 11
AW471332
LOCUS
DEFINITION
 494 bp mRNA linear EST 24-FEB-2000
 xw57hl1.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2832165 3'
 similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
 AW471332.1 GI:7041438
 EST.
 Homo sapiens (human)
ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 1 (bases 1 to 494)
AUTHORS
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
JOURNAL
 Unpublished (1997)
CONTACT
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-x@mail.nih.gov
 Life Technologies catalog #: 11548-013
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 417.

```

11548-013"

Query Match      7.9%; Score 236.8; DB 10; Length 494;
Best Local Similarity 84.7%; Pred. No. 3.4e-42;
Matches 265; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY      315  CCGGGTCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGCGGAGGTGGGTGATC 374
      |||||
Db       181  CCGGGTGTGGTGGCTCATGCTGTAATTCGAGCACATTTGGGAGGCTGAAGTGGGAGATC 240
      |||||

QY      375  ACTTGAGGTCAGGAGTTTCGAGACAGCCTGGCCAAACATGGTGAACCCCTGTCTCTACTAA 434
      |||||
Db       241  ACTTCAGGTCAGGAGTTTCGAGACAGCCTGGCCAAATATGGTGAACCCCGTCTCCACTAA 300
      |||||

QY      435  AATATATAAAAATTAGCTGGGCATGTGGTTCGTACCTGTGTAATCCGATACACTTGGGAGGT 494
      |||||
Db       301  AATATACAAAAATTAGCTGGGTGTGGTAGAGTGTGCCCTGTAATCCCGACTACTTTGGGAGGC 360
      |||||

QY      495  TGAGGCAGGAGAAATCGCTTCAACCCGGGAGCGAGATGTTCCAGTGAACCAAGACTGTGCC 554
      |||||
Db       361  TGAGGCAGGAGAAATCGTTCGAACTGGGAGCGAGAGTTCCAGTGAACCAAGATTGGCCC 420
      |||||

QY      555  ACTGCATCTCCAGTCTGGGCAACAGAGTGAGCCCTCCATCTCAAAACAAACAAAAAGCA 614
      |||||
Db       421  ACTGCATCTCAGCCTGGGCAACAGAGCGAGACTCCATCTCANAAAAAAGTGAAATGCA 480
      |||||

QY      615  GTGCCCATCATGT 627
      |||||
Db       481  CGTGCCTTCTTAT 493

RESULT 12
BQ880197/c
LOCUS      BQ880197      932 bp      mRNA      linear      EST 16-AUG-2002
DEFINITION AGENCOURT 8122269 lupski dorsal root ganglion Homo sapiens cDNA
            clone IMAGE:6179509 5', mRNA sequence.
ACCESSION  BQ880197
VERSION    BQ880197.1 GI:22272205
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 932)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLMI3560 row: m column: 14
            High quality sequence stop: 604.
            Location/Qualifiers
                1..932
FEATURES             source

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6179509"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="lupski_dorsal_root_ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:

```

5'-TCGACCCACGCTCCG-3' and
5'-GACATGTTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupeki, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Query Match 7.9%; Score 236.8; DB 13; Length 932;
Best Local Similarity 63.4%; Pred. No. 3.8e-42;
Matches 378; Conservative 0; Mismatches 217; Indels 1; Gaps 1;
QY 2407 GCTCAGCTTCCCATCTTAGCAGCTTTGGAGGCTGAGTGGGTGATTAACCTCAGTCAG 2466
Db GCTCAGGCTTATAATCCCGAGCAGCTTTGGAGGCTGAGGAGATGATCACCCTGGTGG 632
QY 2467 GAGTTCAAGACGAGCTAGACAAACGTAGAGAAACCCCATCTCTACTGAGATA-TAAAT 2525
Db GAGTTGAGACCATCTCGCCACACATGCGAAACCCCGTCTCTACTGAAATACAAAT 572
QY 2526 TAGCTGGCGTAGTGGCATATGCTGTAAATCCAGCTAGTCGGGAAGCTGAGGAGGAGA 2585
Db TAGCCAGCGTGGTGGCAGGCGCCCTGTAGTCCAGCTACTTTGGGAGGCTGAGGAGGAGA 512
QY 2586 ATCGCTTGAACCCGAGAGCGAGGTTGGGTGAGCTGAGTCGTGCGCATTCACCTCCAG 2645
Db ATCACTTGAACCCGGGAGGAGAGTTGAGTGGAGCGGAGATCGTGCCACTGCATCCAG 452
QY 2646 CTTGGGTGACAAAGCAGACCTCGTCTCAATAATAATAATATACAAACAGAGAGGAGC 2705
Db CTTGGGAGCAGAGCGAGACTCGCTCCAGAAACCAAAAGAACTAATAAATATTG 392
QY 2706 CTGGGTGATCCAGCTACTACTTTTTCAGGAGATGTACTCCCTTACCCAGGCGAAGG 2765
Db GGAAACCTTCTTGTACTAATACATACCTAAACCTTTAATCTAGAACAGATGCCTACATC 332
QY 2766 ATGGGAGAACAGTTTGTATTATGATTTATTGAGCAGCTACTGAGTCTCATCCCTGGG 2825
Db CAGTATTAAATAGGTAATAATCTAACCAAGTATATCTTTTATTACCTACCACTTTA 272
QY 2826 TAGGTGGAATGGAATCAGATGGAGCTTGAAGAGTCCCTCAGGGAACCTCAGTAGAAA 2885
Db CTAAATAGTTGAAAAAAGCGCAAGACAGAGGCGATCATCTCTCCAGTGATATAA 212
QY 2886 GAAGAGGAATCGCGCGGCGGTGCTCAGCCTGTAAATCCCAACACTTTGGAGGCT 2945
Db AGAAGCATGATGGGCTGGCGGCTGAGCTCAAGCTGTAAATCCAGCACTTTGGGAAGCC 152
QY 2946 GAGGTGGGTGATCACAAGGTGAGGAGATCGAGACCATCTGGCTAAACACAGTGAA 3001
Db AAGGCTGCAGATCAGCAAGTCAGAGATCGAGACCATCTGGCCCAACATGTTGAA 96

RESULT 13
A0045514
LOCUS A0045514 498 bp DNA linear GSS 14-APR-1999
DEFINITION RPC111-33E7.TJ RPC1-11 Homo sapiens genomic clone RPC1-11-33E7,
genomic survey sequence.
ACCESSION A0045514
VERSION A0045514.1 GI:3314441
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 498)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPC1-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Class: BAC ends.

FEATURES
source

1..498
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7512390"
/db_xref="taxon:9606"
/clone="RPC1-11-33E7"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPC1-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"

ORIGIN

Query Match 7.9%; Score 236; DB 28; Length 498;
Best Local Similarity 84.2%; Pred. No. 5.1e-42;
Matches 266; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 236 TGATAACAGTGCCCATCAGCGGGTGAGTGGCTAGTCCCTGAAATCCCAACACTTTGGG 355
Db TTATGATAGTATTCCTCAGCTGGGCGGTGGTCAACGCTGTAATCCAGCAGCTTTGGG 196
QY 356 AGGCGGAGTGGGTGGATCCTTGAGGTGAGGAGTTCGAGACCAAGCTGGCCCAACATGTT 415
Db AGGCCAAGTGGAGTGATCCTGAGGTGAGGAGTTCGAGACCAAGCTGGCCCAACATGTT 256
QY 416 GAAACCTGTCTCTACTAAAAATATAAAAAATAGCTGGGCGATGGTGGTACCTGTAA 475
Db GAAACCTGTTTGTACTAAAAATACAAAAATTAACCTGGGCGATGGTGGGCACTATAA 316
QY 476 TCCAGATACTTGGGAGGTTCAGGCGAGAGATCGCTTGAACCCGGGAGGAGATGTTGC 535
Db TCCAGCTACTTGGGAGGCTGAGGCAAGAGATCACTTGAACCTGGGAGGAGATGTTGC 376
QY 536 AGTGAACCAAGACTGTGCCACTGCACTCCAGTCTGGGCAACAGAGTGAAGCTCCATCTCA 595
Db AGTGAGCGAGATCGCGCATTTGCATCCAGCTAGGCGAGCAAGTGAAGTGAAGTCCATCTCA 436
QY 596 AACAAACAAACAAAAA 611
Db AAAAAAAAAAAAAAAAAA 452

RESULT 14
A0035209
LOCUS A0035209 656 bp DNA linear GSS 01-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-010M15.F, genomic survey sequence.
ACCESSION A0035209
VERSION A0035209.1 GI:16562082
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Tokui,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
JOURNAL Unpublished
COMMENT 2 (bases 1 to 656)

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
 Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou,Tsuri-Ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:chimbese@sc.riken.go.jp, URL:http://bgp.gsc.riken.go.jp/,
 Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.
PRIMERS
 Sequencing: -21M13
LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI
FEATURES Location/Qualifiers
 source 1..656
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-010M15.F"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
 Query Match 7.9%; Score 236; DB 29; Length 656;
 Best Local Similarity 86.7%; Pred.No.5.3e-42;
 Matches 260; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 QY 313 AGCGGGTGCAGTGGCTAGTGCCTGGAATCCCAACACTTTGGGAGCGGAGGTGGGGA 372
 Db 193 AGCAGAGCAGTAGCTCATGCCGTAAATCCAGACACTTTGGGAGCGGAGCGGGA 252
 QY 373 TCACCTGAGTCCAGGAGTTCACAGACCGCTGGCCACATGGTGAAACCCCTGCTCTACT 432
 Db 253 TCACCTGAGTCCAGGAGTTCACAGACCGCTGGCCACATGGTGAAACCTGCTCTACT 312
 QY 433 AAAAATATAAAATTAGCTGGGATGTTGGTGCTACTGTAAATCCAGATACTTGGGAG 492
 Db 313 AAAAATACAAAAATTAGCTGGGATGTTGGTGCTACTGTAAACCGGAGCGAGATCGTG 432
 QY 493 GTTGAGGCGAGGAATCGCTTGAACCGGAGCGAGATGTTGCAGTGAACCAAGACTGTG 552
 Db 373 GCTGAGGCGAGGAATCGCTTGAACCGGAGCGAGATGTTGCAGTGAACCAAGACTGTG 432
 QY 553 CCACCTGACTCCAGTCTGGCCACAGAGTGGAGCTCCATCTCAAAACAAACAAAAG 612
 Db 433 CCACCTGACTCCAGCTGGGCGACTGAGTGAGAATCCATCTCAGTCAATCAACGAGTCAG 492
RESULT 15
 AQ590440 466 bp DNA linear GSS 08-JUN-1999
 HS 5388 Al D12 T7A RPCI-11 Human Male BAC Library Homo sapiens
 Genomic Clone Plate=964 Col=23 Row=G, genomic survey sequence.
ACCESSION AQ590440
VERSION AQ590440.1 GI:5022092
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 466)
AUTHORS Mahalir,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
TITLE Sequencing-tagged connectors: A sequence approach to mapping and
 scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:14:41; Search time 121 Seconds
(without alignments)
1314.225 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLIALALLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_nhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2615	100.0	504	4	Q9BZ66
2	2527	96.6	564	4	Q96SQ9
3	2060.5	78.8	501	11	Q9BEX6
4	1313	50.2	491	6	Q29532
5	1301	49.8	493	6	Q8SQ67
6	1300	49.7	323	11	Q8Q027
7	1298	49.6	491	6	Q29516
8	1296	49.6	491	11	Q9WDD0
9	1261.5	48.2	494	11	Q9QZ50
10	1257	48.0	491	11	Q64584
11	1255	48.0	491	11	Q8QXG6
12	1247	47.7	491	11	Q64463
13	1243.5	47.6	494	11	Q9WV19
14	1242.5	47.5	494	6	Q8SQ68
15	1236	47.3	493	11	Q8QXK1
16	1236	47.3	494	11	Q91X75

17	1233	47.2	494	11	Q91XG2
18	1226	46.9	493	11	Q9R0R6
19	1224	46.8	494	11	Q9JJ02
20	1221.5	46.7	494	11	Q9Z1H8
21	1214.5	46.4	495	11	Q64583
22	1210	46.3	491	11	Q64460
23	1210	46.3	495	11	Q9JW91
24	1199.5	45.9	488	6	Q7YSF7
25	1180.5	45.1	494	13	Q92129
26	1175.5	45.0	494	4	Q13120
27	1168	44.7	494	11	Q8QZM4
28	1166.5	44.6	492	11	Q8VCM9
29	1164	44.5	487	11	Q8CIE7
30	1156	44.2	491	6	Q8WNE1
31	1152.5	44.1	494	6	Q8SQ66
32	1127	43.1	494	13	Q8QFT4
33	1125	43.0	495	4	Q8WJ2
34	1089.5	41.7	490	6	Q8SQ65
35	1085.5	41.5	487	4	Q16756
36	1085	41.5	496	13	Q7ZX81
37	1083.5	41.4	490	11	Q8C7K2
38	1074.5	41.1	490	11	Q91WN9
39	1071.5	41.0	477	4	Q16872
40	1069	40.9	474	11	Q9CVC8
41	1056.5	40.4	490	11	Q9BBD9
42	1052.5	40.2	485	4	Q9UC29
43	1051.5	40.2	490	4	Q8WNB1
44	1046.5	40.0	495	11	Q91Y29
45	1043.5	39.9	490	11	Q91X77

ALIGNMENTS

RESULT 1

Q9BZ66	Q9BZ66	PRELIMINARY;	PRT;	504 AA.
AC	Q9BZ66;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Cytochrome P450 2S1 (Similar to cytochrome P450, subfamily IIS, polypeptide 1).			
DE	CP2S1.			
GN	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21092856; PubMed=11181079;			
RA	Rylander T., Neve E., Ingelman-Sundberg M., Oscarson M.,			
RT	"Identification and Tissue Distribution of the Novel Human Cytochrome P450 2S1 (CYP2S1).";			
RL	Biochem. Biophys. Res. Commun. 281:529-535 (2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood;			
RC	Strausberg R.;			
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.			
DR	EMBL; AF335278; AAK13498.1; -;			
DR	EMBL; BC033691; AAK33691.1; -;			
DR	PIR; JC7613; JC7613.			
DR	HSSP; P00179; 1DT6.			
DR	GO; GO:0005783; C:endoplasmic reticulum; NAS.			
DR	GO; GO:0004497; F:monooxygenase activity; NAS.			
DR	GO; GO:0006118; P:electron transport; NAS.			
DR	InterPro; IPR001128; Cytochrome_P450.			
DR	Pfam; PF00067; P450; 1.			
DR	PRINTS; PR00385; P450.			
DR	PROSITE; PS00086; CYTOCHROME P450; 1.			
KW	Heme; Monooxygenase; Oxidoreductase.			

```
KW Hypothetical protein; Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 564 AA; 62231 MW; 03B2E9376A494D12 CRC64;

Query Match 96.6%; Score 2527; DB 4; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.2e-192;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLLTALSSTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60
DB 1 MEATGTWALLALALLLLLTALSSTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60

QY 61 SKKYGPEVTIYLGPRPVVVLVGQAVREALGQAEFSGRTVAMLEGTDFGHGVFFSN 120
DB 61 SKKYGPEVTIYLGPRPVVVLVGQAVREALGQAEFSGRTVAMLEGTDFGHGVFFSN 120

QY 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOQTEGRPPDPSLLLAQATSN 180
DB 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOQTEGRPPDPSLLLAQATSN 180

QY 181 VVCSLLFLGRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
DB 181 VVCSLLFLGRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPTEFTNKMLMTVIYLL 300
DB 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPTEFTNKMLMTVIYLL 300

QY 301 FAGTMTVSTTVGYTLMLLMKYPHVQKWREELNRELGAGAPSLGDRTRLPTDVLHFA 360
DB 301 FAGTMTVSTTVGYTLMLLMKYPHVQKWREELNRELGAGAPSLGDRTRLPTDVLHFA 360

QY 361 QRLALLVPMGIPTLMTTRFRGYTLPGQTEVPFLGLSILHDPNIFKHPEEFPDRFLDA 420
DB 361 QRLALLVPMGIPTLMTTRFRGYTLPGQTEVPFLGLSILHDPNIFKHPEEFPDRFLDA 420

QY 421 DGRFRKHEAFLPFLSKRVCGLGELAKAELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFLSKRVCGLGELAKAELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480

QY 481 LFNIPPA 487
DB 481 LFNIPPA 487

RESULT 3
Q9DBX6 PRELIMINARY; PRT; 501 AA.
AC Q9DBX6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 1200011C15RIK protein (Similar to cytochrome P450 2B1).
GN CYP2B1 OR 1200011C15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesele G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
```

```
SQ SEQUENCE 504 AA; 55816 MW; 853513370F65E25A CRC64;

Query Match 100.0%; Score 2615; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 1e-199;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLLTALSSTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60
DB 1 MEATGTWALLALALLLLLTALSSTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRL 60

QY 61 SKKYGPEVTIYLGPRPVVVLVGQAVREALGQAEFSGRTVAMLEGTDFGHGVFFSN 120
DB 61 SKKYGPEVTIYLGPRPVVVLVGQAVREALGQAEFSGRTVAMLEGTDFGHGVFFSN 120

QY 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOQTEGRPPDPSLLLAQATSN 180
DB 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOQTEGRPPDPSLLLAQATSN 180

QY 181 VVCSLLFLGRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
DB 181 VVCSLLFLGRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPTEFTNKMLMTVIYLL 300
DB 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPTEFTNKMLMTVIYLL 300

QY 301 FAGTMTVSTTVGYTLMLLMKYPHVQKWREELNRELGAGAPSLGDRTRLPTDVLHFA 360
DB 301 FAGTMTVSTTVGYTLMLLMKYPHVQKWREELNRELGAGAPSLGDRTRLPTDVLHFA 360

QY 361 QRLALLVPMGIPTLMTTRFRGYTLPGQTEVPFLGLSILHDPNIFKHPEEFPDRFLDA 420
DB 361 QRLALLVPMGIPTLMTTRFRGYTLPGQTEVPFLGLSILHDPNIFKHPEEFPDRFLDA 420

QY 421 DGRFRKHEAFLPFLSKRVCGLGELAKAELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFLSKRVCGLGELAKAELFFFTTILQAFSLSPCPDPTLSLKPTVSG 480

QY 481 LFNIPPAQLQVRPTDLHSTTQTR 504
DB 481 LFNIPPAQLQVRPTDLHSTTQTR 504

RESULT 2
Q96SQ9 PRELIMINARY; PRT; 564 AA.
AC Q96SQ9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ14699.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masuko Y., Kanehori K.;
RL "NEDO human cDNA sequencing project."
CC Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DE ENBL; AK027605; BAB55227.1; -.
DR Genew; HGNC:15654; CYP2S1.
DR GO; GO:0004497; P:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
```


RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming D.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AK004699; BAB23484.1; -;
 DR EMBL; AK087069; BAC39794.1; -;
 DR HSSP; P00179; 1DT6.
 DR MGD; MGI:1921384; Cyp2s1.
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; p450; 1.
 DR PRINTS; PR00385; P450.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 501 AA; 55631 MW; D12575F3F8D7B019 CRC64;

Query Match 78.8%; Score 2060.5; DB 11; Length 501;
 Best Local Similarity 78.8%; Pred. No. 1.4e-155;
 Matches 391; Conservative 50; Mismatches 54; Indels 1; Gaps 1;

QY 1 MEATGTTALLLA-LALLLLLTALSGTRAGHLPPGTPPLLLGNLLQRLGALYGLMR 59
 Db 1 MEASTWALLALLLLLSLTFTPTARGYLPFGPTPLLLGNLLQRLGALYGLLR 60
 QY 60 LSKYGVFTYILGPWPVPVVLVGOEAVREALGQAEFEFGRTVAMLEGTFDGHGVFFS 119
 Db 61 LSKYGVFTYILGPWPVPVVLVGHDAVREALGQAEFEFGRTVATLDTKTFDGHGVFFA 120
 QY 120 NGERWQLRFTMLALDLGMRKEGELIQAEARCLVETPQGTGRTFPDPSLLLAQTS 179
 Db 121 NGERWQLRFTMLALDLGMRKEGELIQAEVQSLVEAFQKTEGRTFPDPSLLLAQTS 180
 QY 180 NVVCSLLFGLRFSYEDKEFOAVRAAGTLLGVSSQGGQYEMFSLRPLRPGHKLH 239
 Db 181 NVVCSLVFGLRPLVDKFEQAVTQAASGTLGTSPPWQAYEMFSLRPLRPGHKLH 240
 QY 240 HVSTLAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEEQNGTFTNKNMLMTVYL 299
 Db 241 HLGTLAFTVQVQHQGNLDASGPARDLVDAFLKMAQEQNGTFTNKNMLMTVYL 300
 QY 300 LFGATMTVSTVGYTLTLLMKYPHVQKWRVREELNRELGAQPSLGRDTRLPYTDVLHE 359
 Db 301 LFGATMTVSTVGYTLTLLMKYPHVQKWRVREELNRELGAQPSLGRDTRLPYTDVLHE 360
 QY 360 AQLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLLGSLIHDPNIPKHPPEFNPRFLD 419
 Db 361 AQLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLLGSLIHDPNIPKHPPEFNPRFLD 420
 QY 420 AGRFRKHEAFLPFLSKRYCLCEGLAKAELEFPTTILQAFSLSPCPDPTLSLKPVS 479
 Db 421 EDGLRKHAEFLPFLSKRYCLCEGLAKAELEFPTTILQAFSLSPCPDPTLSLKPVS 480
 QY 480 GFENIPPAFOLQVRPT 495
 Db 481 GFENIPDFQLRWPT 496

RESULT 4
 Q29532

ID Q29532 PRELIMINARY; PRT; 491 AA.
 AC Q29532; Q29515;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Cytochrome P450 2B-BX (BC 1.14.14.1).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93349061; PubMed=8346920;
 RA Ryan R., Grimm S.W., Kedzie K.M., Halpert J.R., Philpot R.M.;
 RT "Cloning, sequencing, and functional studies of phenobarbital-
 RT inducible forms of cytochrome P450 2B and 4B expressed in rabbit
 RT kidney";
 RL Arch. Biochem. Biophys. 304:454-463 (1993).
 RN [2]
 RP SEQUENCE OF 2-491 FROM N.A.
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Kidney;
 RA Ryan R., Grimm S.W., Kedzie K.M., Halpert J.R., Philpot R.M.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; S64259; AAB27705.1; -;
 DR EMBL; L10912; AAA31432.1; -;
 DR PIR; S35666; S35666.
 DR HSSP; P00179; 1DT6.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR008068; EP450_CYP2B.
 DR Pfam; PF00067; p450; 1.
 DR PRINTS; PR01685; EP450ICYP2B.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Electron transport; Heme; Membrane; Monooxygenase; Oxidoreductase.
 FT METAL 436 436 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 491 AA; 55732 MW; E3929C649FC66F17 CRC64;

Query Match 50.2%; Score 1313; DB 6; Length 491;
 Best Local Similarity 50.7%; Pred. No. 4.7e-96;
 Matches 24; Conservative 99; Mismatches 137; Indels 4; Gaps 3;

QY 7 WALLALALL--LILLTLALSGTRAGHLPPGTPPLLLGNLLQRLGALYGLMRLSKY 64
 Db 3 FSLILLALFAGLILLLLFRGHFKAGRLPPGTPPLVGLNLLQMDRKLRSFRLREKY 62
 QY 65 GPVFTYILGPWPVPVVLVGOEAVREALGQAEFEFGRTVAMLEGTFDGHGVFFSNGERW 124
 Db 63 GDVFTVYLGSL-RPWVVLGCTDAIRALVDQAEAFSGRGKIAVVDPIFGYGVIFANGERW 121
 QY 125 RQLRKFTMLALDLGMRKEGELIQAEARCLVETPQGTGRTFPDPSLLLAQTSNVVCS 184
 Db 122 RAURRFSRLTRDFMGMRKSVREIRIQEAEARCLVEELRKSQALLDNTLLFHSITSNICS 181
 QY 185 LLFGLRFSYEDKEFOAVRAAGTLLGVSSQGGQYEMFSLRPLRPGHKLHVVSTL 244
 Db 182 IVFGKFDYKDFVFLRLDLLDFQSFSLISFSQVFLSGFLKHPFGTHQIRNLQEI 241
 QY 245 AAFVTRVQVQHQGNLDASGPARDLVDAFLKMAQEEQNGTFTNKNMLMTVYLLFAGT 304
 Db 242 NTFITQVSKHRATLPSNP-RDFIDVYLLRMEKDSDFSSEFHQNLILTVLSLFAQT 300
 QY 305 MTVSTVGYTLTLLMKYPHVQKWRVREELNRELGAQPSLGRDTRLPYTDVLHEAQRLL 364
 Db 301 ETTSTTLRYGFLMLKYPHTVTRVQKEIEQVIGSHRPPALDDRAKMPYTDVAVHEIQRLG 360
 QY 365 ALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLLGSLIHDPNIPKHPPEFNPRFLDADGRF 424
 Db 361 DLIPFGVPHIVTDQFRGVYIPKNTVEFVLSSALHDPYFKTPTNPFEGHFDANGAL 420

Db	362	IPMGVPHVTVDKTHFRGYLLPKGTEVYDILTSALHDPFRYFENPEAFNPDPHFIDATGALKK	421
Qy	427	HEAPFLPSLGHKRVCLGEGIAKAEPLFFTTILQAFSLSPCPDPTLSLKPTVSGLFNIPP	486
Db	422	NDAFMPSIGIKRICLGEGIAKAEPLFFTTILQAFSLSPCPDPTLSLKPTVSGLFNIPP	481
Qy	487	AFOLQVRP	494
Db	482	PYQIQFLP	489
RESULT 6			
Q8K027	PRELIMINARY; PRT; 323 AA.		
ID	Q8K027		
AC	Q8K027;		
DT	01-OCT-2002 (TremBLrel. 22, Created)		
DT	01-OCT-2002 (TremBLrel. 22, Last sequence update)		
DT	01-JUN-2003 (TremBLrel. 24, Last annotation update)		
DE	Similar to RIKEN cDNA 1200011C15 gene (Fragment).		
GN	CYP2S1 OR 1200011C15RIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Colon;		
RA	Strausberg R.;		
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.		
CC	-!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.		
DR	EMBL; AB034202; AAH34202.1; -.		
DR	MGD; MG1:1921384; Cyp2s1.		
DR	GO; GO:0004497; F:monooxygenase activity; IEA.		
DR	GO; GO:0006119; P:electron transport; IEA.		
DR	InterPro; IPR001128; Cytochrome_P450.		
DR	Pfam; PF00067; P450; 1.		
DR	PRINTS; PR00385; P450.		
KW	Heme; Monooxygenase; Oxidoreductase.		
FT	NON TER		
SQ	SEQUENCE 323 AA; 36065 MW; 71BD0875E35C0E48 CRC64;		
Query Match 49.7%; Score 1300; DB 11; Length 323;			
Best Local Similarity 76.1%; Pred. No. 2.8e-95;			
Matches 242; Conservative 37; Mismatches 39; Indels 0; Gaps 0;			
Qy	178	TSNVVCSLLFGHRSYEDKEFOAVRAAGTLLGVSSQGGQYEMFSWFLRPLPGPHKQL	237
Db	1	TSNVVCSLVFGIRLPYDDKEFOAVIQAAAGTLLGISPMWQAYEMFSWLLQPLPGPHQL	60
Qy	238	LHHVSTLAAFTVRQVQOHGNDASGPARDLVDAFLKKMAQBEONPGTEFTNKNMLMTVI	297
Db	61	QHHLGTAAFTTQQVQKHQGRFQTSGPARDVDVDAFLKKMAQEQDQDGTETFEKNLLMTVT	120
Qy	298	YLLFAGTMTVSTTVGYTLLLMKYPHQKVRRELNRELGAGQAPSLGDRTRLPYTDVLA	357
Db	121	YLLFAGTMTIGATIRVALLLLRYPQVQRVREELIQELGPGRAPSLSDRVRLPYTDVLA	180
Qy	358	HEAQRLLALVPMGIPRTLMRTTRFRGYTLPQGTVEVFLIGLSILHDNIFKHPEEFPDRF	417
Db	181	HEAQRLLALVPMGMPHTITRTTCFRGYTLPKGTVEVFLIGLSILHDPAVFQPGFHPGRF	240
Qy	418	LDADGFRKHEAPLPESLGRVCLGSLAKAEPLFPETTLQAFSLSPCPDPTLSLKPT	477
Db	241	LPEDGRLRKHEAPLPYSLGRVCLGSLAKAEPLFPETTLQAFSLSPCPDPTLSLKPA	300
Qy	478	VSGLFNIPPAFQLQVRPT	495
Db	301	ISGLFNIPDFQLRVWPT	318
RESULT 7			
Q29516			

QY	425	RKHEAFLPSPGKRVCLGEGIAKAEPLFFFTTILQAFSLSPCPDPTTSLKPTVSGLFNI	484
DB	421	KRNEGMPSPGKRVCLGEGIAKAEPLFFFTTILQAFSLSPCPDPTTSLKPTVSGLVN	480
QY	485	PPAFOLQ 491	
DB	481	PPSYQIR 487	
RESULT 5			
Q8SQ67	PRELIMINARY; PRT; 493 AA.		
ID	Q8SQ67		
AC	Q8SQ67;		
DT	01-JUN-2002 (TremBLrel. 21, Created)		
DT	01-JUN-2002 (TremBLrel. 21, Last sequence update)		
DT	01-JUN-2003 (TremBLrel. 24, Last annotation update)		
DE	Cytochrome P450 2B22.		
GN	CYP2B22.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RA	Kojima M.;		
RL	"Cloning of pig cytochrome P450 cDNAs and their expressions in tissues";		
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.		
CC	-!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.		
DR	EMBL; AB052256; BAB85662.1; -.		
DR	GO; GO:0004497; F:monooxygenase activity; IEA.		
DR	GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.		
DR	GO; GO:0006118; P:electron transport; IEA.		
DR	InterPro; IPR001128; Cytochrome P450.		
DR	InterPro; IPR008088; EP450_CYP2B.		
DR	Pfam; PF00067; P450; 1.		
DR	PRINTS; PR01685; EP450ICYP2B.		
DR	PRINTS; PR00385; P450.		
DR	PROSITE; PS00086; CYTOCHROME P450; 1.		
KW	Heme; Monooxygenase; Oxidoreductase.		
SQ	SEQUENCE 493 AA; 56154 MW; 7F6322C41E4077E1 CRC64;		
Query Match 49.8%; Score 1301; DB 6; Length 493;			
Best Local Similarity 49.2%; Pred. No. 4.2e-95;			
Matches 240; Conservative 104; Mismatches 142; Indels 2; Gaps 2			
QY	7	WALLALALLLLLTALSGTRAGHLLPPGTPPLPLGNLLQLRPGALYSGLMRLSKYKGP	66
DB	4	FSAILLLVIGVFWLLMGHSPSSQGLPPGPRPLFLGNLQMDRKGLKSFQALRERYGD	63
QY	67	VFTYILGPMPVTVVGOEAVREALGQAEFFSGRGTVAMLEGTFDHGCVFFSNGERWRQ	126
DB	64	VFTVILGP-RPVNMI CGTEAIRALVDQAEAFSGRKIAVVEPIFGYGVIFSGERWKT	122
QY	127	LRKFTMLALRDLGMGRGEBELIQAEARCLVETFGQTEGRPDPSLLLLAQATSNVVCSSL	186
DB	123	LRRFSLATMRD FGLGKFSVEERIQEAEACLVEELRKSQALQDPTFFYFHSSTANIICSV	182
QY	187	FGLRFSYEDKEFOAVRAAGTLLGVSSQGGQYEMFSWFLRPLPGPHKQLLHHVSTLAA	246
DB	183	FGKRFAVRDPEFLQLDLFLQTFLTISSPFSQSLFELYSAPLKYFPGSHRQIYKNLQQVNA	242
QY	247	FTVROVQOHGNDASGPARDLVDAFLKKMAQEQNPGTETFNKNMLMTVIVILLFAGTMT	306
DB	243	FIGSVEKHRETLDPSP-RDLTDSYLLRMDKESNPSEFFHQNLILTVLSLFPAGTET	301
QY	307	VSTTVGYTLLLMKYPHQKVRBELNRELGAGQAPSLGDRTRLPYTDVILHEAQRLLAL	366
DB	302	TSTTLRYGFLMLKYPHITRIHKIEIDRVIGSHRPPALDDRAKMPYMDAVIHEMQRFGL	361
QY	367	VPMGIPRTLMRTTRFRGYTLPQGTVEVFLIGLSILHDNIFKHPEEFPDRFLDADGRFRK	426

Qy	309	TTVGTYLLLMKYPHVQKWREELNRELGAGQARSLGDRTRLPYTDVHLHQAORLLALVP	368
Db	305	TTLRGYFLMLKYPHVAEKVQKEIDQVIGSHRLPTLDDRTKMPYTDVAIHEIQRFSDLP	364
Qy	369	MGIPRLMRTTRFRGYTLPQGTVEVPLGLSLHLDNIFKHPEEFPNDRFLDADGRFKHE	428
Db	365	IGVHRVTKMTFRGYLLPKTEVYPIUSSALHDPQYEQPDSFPNDRFLDANGALKKSE	424
Qy	429	AFLPSLGRKVCGLGSLAKAEFLFFTTILQAFSLSPCPPDTLSIKPTVSLGNIPPAF	488
Db	425	AFLPSTGKRLGSLIARNELFLFTLSILQNFVASHVAPKDIDLTPKESGIGKIPPT	484
Qy	489	QL 490	
Db	485	QI 486	
RESULT 9			
ID	Q9QZ50	PRELIMINARY;	PRT; 494 AA.
AC	Q9QZ50;		
DT	01-MAY-2000 (TremBLrel. 13, Created)		
DT	01-MAY-2000 (TremBLrel. 13, Last sequence update)		
DT	01-JUN-2003 (TremBLrel. 24, Last annotation update)		
DE	Cytochrome P-450.		
OS	Mesocricetus auratus (Golden hamster).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;		
OC	Mesocricetus.		
OX	NCBI_TaxID=10036;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RX	MEDLINE=99453747; PubMed=10524222;		
RA	Tohkin M., Kurose K., Isezaki E., Fukuhara M.;		
RT	"Molecular cloning, heterologous expression, and characterization of a novel member of CYP2A in the Syrian hamster."		
RL	Biochim. Biophys. Acta 1446:438-442 (1999).		
CC	- I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.		
DR	EMBL; D86952; BAA85463.1; -		
DR	HSSP; P00179; 1DT6.		
DR	GO; GO:0004497; F:monooxygenase activity; IEA.		
DR	GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . . ; IEA.		
DR	GO; GO:0006118; P:electron transport; IEA.		
DR	InterPro; IPR001128; Cytochrome P450.		
DR	InterPro; IPR008067; EP450_CYP2A.		
DR	Pfam; PF00067; P450; 1.		
DR	PRINTS; PR01684; EP450ICYP2A.		
DR	PRINTS; PR00385; P450.		
DR	PROSITE; PS00086; CYTOCHROME P450; 1.		
KW	Heme; Monooxygenase; Oxidoreductase.		
SQ	SEQUENCE 494 AA; 56544 MW; A0A37193E2BC5849 CRC64;		
Query Match			
Best Local Similarity 48.2%; Score 1261.5; DB 11; Length 494;			
Matches 247; Conservative 92; Mismatches 146; Indels 9; Gaps 4;			
Qy	1	MEATGWALLALALLLLTALSQT----RARGHLPPGPTPLPLGNLLQLRPGALYSG	56
Db	1	MLASG---LLLVTVLAVLVLLMSVWKQKLSGLPKPGTPLPFGNYLQLNTEQMYNS	57
Qy	57	LMRLSKYGPVFTIYLGFWRPVVVLVQGBAVREALGQAEFEFSGRGTVAMLEGTFDGHV	116
Db	58	LMKISIRYGPVFTIHLGP-RPIVVLVQGBAEVAVQAEFEFSGRGQATFDMLFKQYGV	116
Qy	117	FPNGEWRQLRKFTMLALRDLMGKREGBELIAEACRIVTFQGTGRRFPFPPSLLLAQ	176
Db	117	AFSSGEARAKRRFSIATLRDFGVGKGIIEIRIQEEAGFLLEAFKTKNGALIDTFYLSR	176
Qy	177	ATSNVCSLLFLGRFSYEDKEFOAVVRAAGTLLGVSSQGGQTVEMFSLRPLPGPHKQ	236
Db	177	TVSNVISSIVFGDRFDEKELSLLRMLLGSFQGTGTSTGQLYEMFSSVMKHLPGPQQQ	236
Query Match			
Best Local Similarity 48.1%; Score 1257; DB 11; Length 491;			
Matches 235; Conservative 100; Mismatches 145; Indels 2; Gaps 2;			
Qy	9	LLIALALLLLTALSCTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRLSKYGPVF	68
Db	7	LLIALLVGFLLLVVRGHPKSGNFPFGPRLPLGNLLQLDRGGLLNSFMQLREKYGVDF	66
Qy	69	TIYLGWRPVVVVLVQGBAVREALGQAEFEFSGRGTVAMLEGTFDGHVFSNGERWQLR	128
Query Match			
Best Local Similarity 48.8%; Pred. No. 1.3e-91;			
Matches 235; Conservative 100; Mismatches 145; Indels 2; Gaps 2;			
Qy	9	LLIALALLLLTALSCTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRLSKYGPVF	68
Db	7	LLIALLVGFLLLVVRGHPKSGNFPFGPRLPLGNLLQLDRGGLLNSFMQLREKYGVDF	66
Qy	69	TIYLGWRPVVVVLVQGBAVREALGQAEFEFSGRGTVAMLEGTFDGHVFSNGERWQLR	128

```

Db 67 TVHLGP-RPVNMLCGTDTIKALVGPEDFSGRGIAVIEPIKEXGVIFANGERKALR 125
QY 129 KFTMLALRDLGMKREGBELIOAEARCLVETTCGTGTRPFDPSLLLAQAATSNVVCSSLFG 188
Db 126 RFLATMRDGMGKRSVEERIQEEAQCLVEELRKSQAGPLDPTFLQCITANICISIVFG 185
QY 189 LRFSEYDEKFOAVRAAGCTLLGVSSQGGOTYEMESWFLRPLPGPHKQLLHVSTLAAPT 248
Db 186 BRFDYTDQFLRLLELFFYRFSLLSSFSQVEFFSGFLKYPFGAHRQISKNLQELDIYI 245
QY 249 VRQVOHQGNLDASGARDLVDAFLKMAQEONFGCTFTNKNMLMTVIYLLFAGTMTVS 308
Db 246 GHIVEKRAITLPSAP-RDFIDTYLLRMEKEKSNHTEHHEHNLMLSLSLFFAGTETSS 304
QY 309 TTGVGTYLLMLKYPHVQKWRVRELNRELGAGAPSLGDRTRLPYTDVLAHEAQRLLALVP 368
Db 305 TLLRYGFLMLKYPHVAEKVQREIDHVGISDRPPSLDDRRPKMPYTEAVIYEIQFSDLVP 364
QY 369 MGIPRTLMRTFRGYTLPOGTEVPPLGSLIHDENIFKHPEEFPDRFLDADGRFKHE 428
Db 365 IGPHRVTKDMWERYLLPKTEVYPIRSSALHDPQYFDHDSFNPHEFLDVGALKKSE 424
QY 429 AFLPFSGLKRVCLGEGLAKEAELFFFTTILQAFSLESPPDPTLSLKTPTVSGLEFNIPAF 488
Db 425 AFMPFSTGKICLGEIARNELEFFFTTILQNFSSVSHLAPKDIIDLTPKESIGIKIPY 484
QY 489 QL 490
Db 485 QI 486

```

```

RESULT 11
Q80XG6 PRELIMINARY; PRT; 491 AA.
AC Q80XG6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome P450, family 2 subfamily B polypeptide 29.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Harderian gland;
RA Alvarez J., Vinas J., Dominguez P.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY250712; AAP04405.1; -.
DR GO; GO:0016712; P:oxidoreductase activity, acting on paired d. .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01685; EP450ICYP2B.
DR PROSITE; PS00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 491 AA; 56298 MW; 34AD1B044DC45850 CRC64;

```

```

Query Match 48.0%; Score 1255; DB 11; Length 491;
Best Local Similarity 49.8%; Pred. No. 1.9e-91;
Matches 240; Conservative 91; Mismatches 149; Indels 2; Gaps 2;
QY 9 LLLALLLLLTLALSTRAGHLPQGPTEPLPLGNLLQLRPGALYSGLMRLSKKYGVPV 68
Db 7 LLLTLLLSFLVLLVGVYKTRGHLPGPRPLPLGNLLQLDRGLLNSFRFQEKYGVDF 66
QY 69 TIVLGPWRPVVVLVQEAENREALGGQAEPSGRGTVMLEGTDTGFGVFPFSGNRWRQR 128
Db 67 TLHLGS-KPVVMVLYGTEATREALVDQAEAFSGRTTIAVLKPTMDYGVIAANGRWKTLR 125

```

```

QY 129 KFTMLALRDLGMKREGBELIOAEARCLVETTCGTGTRPFDPSLLLAQAATSNVVCSSLFG 188
Db 126 RFLATMRDGMGKRSVEERIQEEAQCLVEELRKSQAGPLDPTFLQCITANICISIVFG 185
QY 189 LRFSEYDEKFOAVRAAGCTLLGVSSQGGOTYEMESWFLRPLPGPHKQLLHVSTLAAPT 248
Db 186 BRFDYTDQFLRLLELFFYRFSLLSSFSQVEFFSGFLKYPFGAHRQISKNLQELDIYI 245
QY 249 VRQVOHQGNLDASGARDLVDAFLKMAQEONFGCTFTNKNMLMTVIYLLFAGTMTVS 308
Db 246 GHIVEKRAITLPSAP-RDFIDTYLLRMEKEKSNHTEHHEHNLMLSLSLFFAGTETSS 304
QY 309 TTGVGTYLLMLKYPHVQKWRVRELNRELGAGAPSLGDRTRLPYTDVLAHEAQRLLALVP 368
Db 305 TLLRYGFLMLKYPHVAEKVQREIDHVGISDRPPSLDDRRPKMPYTEAVIYEIQFSDLVP 364
QY 369 MGIPRTLMRTFRGYTLPOGTEVPPLGSLIHDENIFKHPEEFPDRFLDADGRFKHE 428
Db 365 IGPHRVTKDMWERYLLPKTEVYPIRSSALHDPQYFDHDSFNPHEFLDVGALKKSE 424
QY 429 AFLPFSGLKRVCLGEGLAKEAELFFFTTILQAFSLESPPDPTLSLKTPTVSGLEFNIPAF 488
Db 425 AFMPFSTGKICLGEIARNELEFFFTTILQNFSSVSHLAPKDIIDLTPKESIGIKIPY 484
QY 489 QL 490
Db 485 QI 486

```

```

RESULT 12
Q64463 PRELIMINARY; PRT; 491 AA.
ID Q64463
AC Q64463;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Testosterone 16a-hydroxylase type a.
GN CYP2B9 OR 16AOH-A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91146586; PubMed=1997326;
RA Lakso M., Masaki R., Noshiro M., Negishi M.;
RT "Structures and characterization of sex-specific mouse cytochrome P-
RT 450 genes as members within a large family. Duplication boundary and
RT evolution."
RL Eur. J. Biochem. 195;477-486(1991).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; M60273; AAA03648.1; -.
DR EMBL; M60267; AAA03648.1; JOINED.
DR EMBL; M60268; AAA03648.1; JOINED.
DR EMBL; M60269; AAA03648.1; JOINED.
DR EMBL; M60270; AAA03648.1; JOINED.
DR EMBL; M60271; AAA03648.1; JOINED.
DR EMBL; M60272; AAA03648.1; JOINED.
DR PIR; I84735; I84735.
DR HGSP; P00179; I076.
DR MGD; MGI:88600; Cyp2b9.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01685; EP450ICYP2B.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 491 AA; 55740 MW; C1E790A7DD1A7298 CRC64;

```

Query Match 47.7%; Score 1247; DB 11; Length 491;
 Best Local Similarity 48.8%; Pred. No. 8.4e-91;
 Matches 235; Conservative 99; Mismatches 146; Indels 2; Gaps 2;

QY 9 LLLALLLLTLLSSTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRLSKK 68
 DB 7 LLLAVLLSLFLLVGRHAKIHGLPPGPHPLPLGNNLQMDRGKLLKCFIQLOEKHGDF 66

QY 69 TIYLGPMRPVVVLVQGEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSNGERWQLR 128
 DB 67 TVHLGP-RPVVVLCTGTITREALVDHAEAFSGRTIAAAQLVMQDYGIFFASGQRWKLIR 125

QY 129 KFTMALRDGLMGKEGEBELIQABARCLVETFGTEGRFPDPSLLLAQATSNVCSLLFG 188
 DB 126 RFLSATMKEFGMGKRSVEERIKKEEAQCLVEELKKGQVPLDPTFLFQCITANILCSIVFG 185

QY 169 LRFSEYEDKEFOAVRAAGTLLGVSSOGGTYEMFSWFLRPLPGPHKQLLHHVSTIAAFT 248
 DB 186 ERFDTDDQFLHLNLMYKIFSLLSFSFGOMELFSGFLKYPFVHRQIVKKQBELLDYI 245

QY 249 VRQVOQHGNLDASGPARDLVDAFLKMAQESONPTEFTNQMLMTVIYLLFAGTMVTS 308
 DB 246 AHSVEKHATLPSAP-RDYIDTYLLRMEKEKSNHNTFHHQNLMMSVLSLFFAGTETTS 304

QY 309 TTVGYTLLMLKYPHVQKVRRELNRELGAQPSLGDRTLPYTDVAVLHEAQRLLALVP 368
 DB 305 ATLHYGVLLMLKYPHVTEKQVEIDQVIGSHRLPTLDDRTKMPYTDVAVLHEIQRSDLVP 364

QY 369 MGIPRLMTTRPRGYTLPGTEVFPVLLGSIILHDPNIFKHPEEFNPDRLDADGRFKHE 428
 DB 365 IGLPHKVIKDTLFRGYLLPKNTEVYFVLSALHDPQYFQPKFNPEHFLDANGALKKCE 424

QY 429 AFLPFLSGKRVCLGEGELAKAELFFFTTILQAFSLSPCPDPLSLKPTVSGLFNPPAF 488
 DB 425 AFLPFTGKRICLGESIAELNELLFFFTTILQAFSLSPVAPKDIDLTPKESGIGKIPPAH 484

QY 489 QL 490
 DB 485 QI 486

RESULT 13
 Q9WV19 PRELIMINARY; PRT; 494 AA.

ID Q9WV19
 AC Q9WV19
 DT 01-NOV-1999 (T-EMBLrel. 12, Created)
 DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Olfactory-specific steroid hydroxylase.
 GN CYP2G1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Olfactory neuroepithelium;
 RX MEDLINE=97288345; PubMed=9143323;
 RA Hua Z., Zhang Q.Y., Su T., Lipinskas T.W., Ding X.;
 RT "cDNA cloning, heterologous expression, and characterization of mouse
 CYP2G1, an olfactory-specific steroid hydroxylase.";
 RL Arch. Biochem. Biophys. 340:208-214(1997).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; L81171; AAD45178.1; -;
 DR HSP; P00179; 1DT6.
 DR MGD; MGI:109612; Cyp2g1.
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01684; EP450ICYP2A.

DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 494 AA; 56819 MW; 63D83B09DB08AB9F CRC64;

Query Match 47.6%; Score 1243.5; DB 11; Length 494;
 Best Local Similarity 48.3%; Pred. No. 1.6e-90;
 Matches 235; Conservative 96; Mismatches 153; Indels 3; Gaps 3;

QY 5 GTWALLLAL-LLLTLLSSTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRLSKK 63
 DB 5 GAFSIFMALCULCSLLILIAWKRTSKGKLLPPGPTPIPLGNFLQVTRTATFOSFQKLQKK 64

QY 64 YGPEVTIYLGPMRPVVVLVQGEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSNGER 123
 DB 65 YGSVTVTVFGP-RPVVVLGCGEAVKEALVDQADDFSGRGEMPTLEKNFGQYGLALSNGR 123

QY 124 WQLAKFTMALRDGLMGKEGEBELIQABARCLVETFGTEGRFPDPSLLLAQATSNVVC 183
 DB 124 WKILRRFSLTVLRNFMGKRSIEERIQEAGYLLBELHVKVGAPIDPTIYLSRTVSNVIC 183

QY 184 SLLFGLRFSYEDKEFOAVRAAGTLLGVSSOGGTYEMFSWFLRPLPGPHKQLLHHVST 243
 DB 184 SVVFGKRPDYQQRFOQLMRMINESFVEMSKPAQLYDMYKVMQYFPGRHVLYNLIED 243

QY 244 LAAFTVRVQVQHQGNLDASGPARDLVDAFLKMAQESONPTEFTNQMLMTVIYLLFAG 303
 DB 244 LKDFIASRVKINEASDPSPN-RDFIDCFILKMQKSDPHTFNLKVLVLTTLNLFEG 302

QY 304 TMTVSTTVGYTLLMLKYPHVQKVRRELNRELGAQPSLGDRTLPYTDVAVLHEAQRLL 363
 DB 303 TETVSSTLYRGYLLLLKYPEVEAKIHEEINQVITGTHRTPRVDDRAKMPYTDVAVLHEIQRL 362

QY 364 LALVPMGIPRTLMRTTRPRGYTLPGTEVFPVLLGSIILHDPNIFKHPEEFNPDRLDADGR 423
 DB 363 TDIVPLGVPHNVTRDTHFRGYLLPKGTDVYPLPGSVLKPQKPYRYDPAFPQHFLEBQGR 422

QY 424 FRKHEAFLPFLSGKRVCLGEGELAKAELFFFTTILQAFSLSPCPDPLSLKPTVSGLFN 483
 DB 423 FKNDFAVVFSSGKRICVGEALARMELFVFTSILQRFSLRSLVPPADIDIAHKISGFNG 482

QY 484 IPPAFQL 490
 DB 483 IPPVVEL 489

RESULT 14
 Q8SQ68 PRELIMINARY; PRT; 494 AA.

ID Q8SQ68
 AC Q8SQ68;
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Cytochrome P450 2A19.
 GN CYP2A19.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RC Kojima M.;
 RT "Cloning of pig cytochrome P450 cDNAs and their expressions in
 tissues";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AB052255; BAB85661.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome P450.
 DR PRINTS; PR008067; EP450_CYP2A.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:14:40 ; Search time 25 Seconds
(without alignments)
1049.734 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1319.5	50.5	494	1	CPG1_RABIT
2	1300	49.7	491	1	CPB4_RABIT
3	1296	49.6	491	1	CPB5_RABIT
4	1296	49.6	491	1	CPB5_RABIT
5	1290	49.3	491	1	CPB1_RAT
6	1280.5	49.0	500	1	CPBA_MOUSE
7	1262	48.3	494	1	CPBB_CANFA
8	1244	47.6	491	1	CPB9_MOUSE
9	1243.5	47.6	494	1	CPG1_RAT
10	1238	47.3	491	1	CPB6_HUMAN
11	1237	47.3	494	1	CPA5_MOUSE
12	1236.5	47.3	492	1	CPB3_MOUSE
13	1235	47.2	494	1	CPA4_MOUSE
14	1234.5	47.2	494	1	CPA3_RAT
15	1214.5	46.4	494	1	CPAD_HUMAN
16	1212.5	46.4	494	1	CPA6_HUMAN
17	1210.5	46.3	492	1	CPBC_RAT
18	1203.5	46.0	494	1	CPAA_RABIT
19	1198.5	45.8	494	1	CPAB_RABIT
20	1199	45.5	494	1	CPAB_RABIT
21	1176.5	45.0	494	1	CPBK_MOUSE
22	1172	44.8	491	1	CPA7_HUMAN
23	1168	44.7	491	1	C2F3_CAPHI
24	1163.5	44.5	494	1	C2F2_MOUSE
25	1160	44.4	491	1	CPCN_RAT
26	1158	44.3	493	1	CPA9_MESAU
27	1150.5	44.0	492	1	CPAC_MOUSE
28	1145	43.8	491	1	C2F4_RAT
29	1138.5	43.5	492	1	CPAL_RAT
30	1108	42.4	491	1	C2F1_HUMAN
31	1107	42.3	491	1	CPH2_CHICK
32	1099.5	42.0	494	1	CPA8_MESAU
33	1098	42.0	487	1	CPCG_RABIT

34	1095.5	41.9	490	1	CPZ4_MOUSE
35	1093	41.8	487	1	CPCS_RABIT
36	1092.5	41.8	490	1	CPC7_RAT
37	1091.5	41.7	492	1	CPA2_RAT
38	1085.5	41.5	490	1	CPC9_HUMAN
39	1084.5	41.5	490	1	CPCJ_HUMAN
40	1083	41.4	491	1	CPHI_CHICK
41	1079.5	41.3	490	1	CPCR_MESAU
42	1078.5	41.2	490	1	CPCA_HUMAN
43	1078.5	41.2	490	1	CPCT_MOUSE
44	1078.5	41.2	490	1	CPZ3_MOUSE
45	1072.5	41.0	490	1	CPCP_MESAU

ALIGNMENTS

RESULT 1

ID	CPG1_RABIT	STANDARD	PRT	494 AA
AC	P24461			
DT	01-MAR-1992 (Rel. 21, Created)			
DT	01-MAR-1992 (Rel. 21, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cytochrome P450 2G1 (EC 1.14.14.1) (CYP11G1) (P450-NMB) (Olfactive).			
GN	CYP2G1.			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
EN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91119408; PubMed=1703755;			
RA	Ding X., Porter T.D., Peng H.M., Coon M.J.;			
RT	"cDNA and derived amino acid sequence of rabbit nasal cytochrome			
RT	P450NMB (P450IIG1), a unique isozyme possibly involved in			
RT	olfaction."			
RL	Arch. Biochem. Biophys. 285:120-125(1991).			
CC	-1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE			
CC	MONOOXYGENASES. THIS ISOZYME SEEMS TO BE IMPLICATED IN OLFACTION.			
CC	-1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +			
CC	oxidized flavoprotein + H(2)O.			
CC	-1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.			
CC	-1- TISSUE SPECIFICITY: Olfactory epithelium.			
CC	-1- SIMILARITY: Belongs to the cytochrome P450 family.			
DR	PIR; S13907; B31944.			
DR	HSSP; P00179; IDT6.			
DR	InterPro; IPR001128; Cytochrome P450.			
DR	InterPro; IPR008067; EP450_CYP2A.			
DR	Pfam; PF00067; P450; 1.			
DR	PRINTS; PR01684; EP450ICYP2A.			
DR	PRINTS; PR00385; P450.			
DR	PROSITE; PS00086; CYTOCHROME P450; 1.			
KW	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;			
KW	Microsome; Endoplasmic reticulum; Olfaction.			
FT	METAL 439 439 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).			
SQ	SEQUENCE 494 AA; 56628 MW; 4DA83F2E1CD230BC CRC64;			

Query Match 50.5%; Score 1319.5; DB 1; Length 494;
Best Local Similarity 50.5%; Pred. No. 8.2e-85;
Matches 248; Conservative 97; Mismatches 143; Indels 3; Gaps 3;

QY	1	MEATGTWALLALAL-LILLTLALSGTRAGHLPQGTPLPLGNLQLRPGALYSGLMR 59
Db	1	MELGGAFTFLALCFSCLLLTAWKVRQVGRPLPGTPIPLGNLQLRQTATFOSFLK 60
QY	60	LSKKGPVPTIYLGWRPVVVLVGQEAVERALGGQAEFSGRGTVMLEGTDFGHGVFFS 119
Db	61	LREKGPVPTVYMGPRPVLVILCGHEAVKEALVDRADEFSGRGELASVERNQGHGVALA 119
QY	120	NGERWQRKFTMLALRDLGMRGEEELIQAEARCLVETFGTGTGPRPDPSSLQAQTS 179
Db	120	NGERWRLRRFSLTILRDFGMGRKSIERIQEAGYLLBEFRKTKGAPIDPTFFLSRTVS 179

DE hydroxylase) (P450-16-alpha) (Clone PF3/46).
GN CYP2B10 OR CYP2B-10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89118235; PubMed=3219345;
RA Noshiro M., Lakso M., Kawajiri K., Negishi M.;
RT "Rip locus: regulation of female-specific isozyme (I-P-450(16 alpha)
RT of testosterone 16 alpha-hydroxylase in mouse liver, chromosome
RT localization, and cloning of P-450 cDNA.";
RL Biochemistry 27:6434-6443(1988).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M21856; AAA40425.1; -;
CC PIR; B31047; B31047.
CC HSP; P00179; LDT6.
CC MGD; MGI:88598; Cyp2b10.
CC InterPro; IPR001128; Cytochrome_P450.
CC InterPro; IPR008068; EP450_CYP2B.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR01685; EP450ICYP2B.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; FALSE NEG.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation.
FT MOD RES 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
FT METAL 445 445 IRON (HEME AXIAL LIGAND).
FT SEQUENCE 500 AA; 56743 MW; F660A0D8D0FBA94 CRC64;
CC
CC Query Match 49.0%; Score 1280.5; DB 1; Length 500;
CC Best Local Similarity 49.7%; Pred. No. 4.4e-82;
CC Matches 244; Conservative 99; Mismatches 137; Indels 11; Gaps 3;
CC
CC QY 9 LLLALLLLLTALSGTRAGHLPPOPTPLPGLNMLLQRLPGALYSGMLSKYGPVF 68
CC DB 7 LLLALLVGFLLLLARGHPKSGNFPFPGPRLPGLNLLQMDRGGLLSLIQREKYGDVF 66
CC
CC QY 69 TIYLGPMRPVVVLVQGAAREALGGQAEFSGTGMLEGTDPDGHGVFNSGERWQLR 128
CC DB 67 TVHLGP-RPVMVLCGTDTREALVGQAEAFSGRTVAVVEPTFKEYGVIFANGEMKTLR 125
CC
CC QY 129 KFTMLALDLGMKREGEELIQAEARCLVETFOGTEGRPDPSPSLLAQAATSNVVCSLLP 188
CC DB 126 RFSLATWRDPMGKRSVEERIQIEAQCVELLRKSQGLDPLDPLFCITANVICSIVFG 185
CC
CC QY 189 LRSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPKQLLHHVSTLAFT 248
CC DB 186 EREYTDQFRLLELFYQVFLSLSSFSQMFELSGFLKYPFGAHRQISKNLQELLDYI 245
CC
CC QY 249 VRQVQHQGNLDASGPARDLVDAFLKMAQEQNPNGTEFTNKNMLMTVILLFAGTIVTS 308
CC DB 246 GHSVERHKATLDPSVP-RDFDIDYLLRMEKEKSNQNAEFHQNLMMVSLSLFFVGTETSS 304
CC 309 TTVGTYTLLMLKYPHVQKVVREELNRELGAQAPSLGDRTRLPYTDVAVLHEAQRLLALVP 368

DB 305 TTLHYGFLMLKYPHVTEKQKEIDQVIGSHRLPTLDRTKMPSYSDAVIHEIQRFSDLP 364
QY 369 MGIPTLMTTRFRPGYLYLPQGTETVFPPLGLSILHDPNIFKHEEENPDRLDADGGRKHE 428
DB 365 IGVHRVTKDTLFRGYLLPKNTEVYPIYSSALHDPQYEQPDSENPDPDLOANGALKSE 424
QY 429 AFLPFP-----SLGKRVCLGEGAKAEFLPFTTILQAFSLAESPCPPDTLSLKPTVS 479
DB 425 AFLPFPSTQGIIFDQKSVGKRICLGESIAESELFLFTSILQNFVASHVAPXDIDITPKES 484
QY 480 GLFNIPPAFQL 490
DB 485 GIGKIPTYQI 495
CC
CC RESULT 7
CC ID_CPB CANFA STANDARD; PRT; 494 AA.
CC AC P24450;
CC DT 01-MAR-1992 (Rel. 21, Created)
CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Cytochrome P450 2B11 (EC 1.14.14.1) (CYP11B1) (P450 PBD-2).
CC CYP2B11.
CC GN Canis familiaris (Dog).
CC OS Canis familiaris (Dog).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
CC OX NCBI_TaxID=9615;
CC RN (1)
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Beagle; TISSUE=Liver;
CC RX MEDLINE=90343348; PubMed=2116765;
CC RA Graves P.E., Elhag G.A., Ciaccio P.J., Bourque D.P., Halpert J.R.;
CC RT "cDNA and deduced amino acid sequences of a dog hepatic cytochrome
CC P45011B responsible for the metabolism of 2',4',4',5',5'-
CC RT hexachlorobiphenyl.";
CC RL Arch. Biochem. Biophys. 281:106-115(1990).
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. THIS ISOZYME SEEMS
CC RESPONSIBLE FOR METABOLISM OF 2',4',4',5',5'-HEXACHLOROBIPHENYL.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By phenobarbital.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M92447; AAA30881.1; -;
CC PIR; S11305; S11305.
CC HSP; P00179; LDT6.
CC InterPro; IPR001128; Cytochrome_P450.
CC InterPro; IPR008068; EP450_CYP2B.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR01685; EP450ICYP2B.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation.
FT MOD RES 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
FT METAL 436 436 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT SEQUENCE 494 AA; 56266 MW; 8E9EA639B4B5E4D3 CRC64;
CC
CC Query Match 48.3%; Score 1262; DB 1; Length 494;
CC Best Local Similarity 49.2%; Pred. No. 8.4e-81;

```

RESULT 8
ID CFB9_MOUSE STANDARD; PRT; 491 AA.
P12730;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B9 [BC 1.14.14.1] (CYP11B9) (Testosterone 16-alpha
DE hydroxylase) [P450-16-alpha] (C1one PF26).
DE CYP2B9 OR CYP2B-9 OR RIP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
TISSUE=Liver;
MEDLINE=89118235; PubMed=3219345;
Noshiro M., Lakso M., Kawajiri K., Negishi M.;
"Rip locus: regulation of female-specific isozyme (I-P-450(16 alpha)
of testosterone 16 alpha-hydroxylase in mouse liver, chromosome
localization, and cloning of P-450 CDNA.";
Biochemistry 27:6434-6443(1988).
-!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
monooxygenases. In liver microsomes, this enzyme is involved in an
NADPH-dependent electron transport pathway. It oxidizes a variety
of structurally unrelated compounds, including steroids, fatty
acids, and xenobiotics.
-!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O.
-!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
-!- SIMILARITY: Belongs to the cytochrome P450 family.
-----
This SWISS-PROT entry is copyright It is produced through a collaboration

```

QY	180	NVVCSSLFGLGFSEYDEKFOAVRVAAAGGTLGVSSGGQTGYEMFWFLRPLPGPHKQLH	233
Dd	160	NVICSVVGKFEFDYDQRFSRLMKMINESVEVMSPWAQLDYMWGVQIYFFGRHNRLYN	239
QY	240	HVSTLAFTVRVOOHOGLNDASGPARDLVDAFLKKWAOEEQNPGTFTNKMLMTVIYL	299
Dd	240	LIEELKDFIASRVKINEASFPDPNP-RDFIDCFLLIKMYQKSDPHEFNKLNLVLTTLNL	298
QY	300	LPAGTMVTSTTVGTVLTLLLMMKYHHVKWKVREELNRELGAQAFLSGDRTRLPTDAVLHE	359
Dd	299	PFAGTETVSSILRYCGFLLLMKYPEVEAKIHHEINQVIGTHRTPRVDRAKMPYTDVAIHE	358
QY	360	AORLALVPMGIPRTLMTTRFRGYTLPOGTEVFPLLSILLHDPNIFKHBEENPDREFLD	419
Dd	359	IQLRLTDIVPLGVPHNVIRDTHFRCGYFLPKGTDVYPLGSVLKDPKFYRYPEAFYPOHFFD	418
QY	420	ADGRERKEAEFLPSLGNKRVCLGSLKAELFLFFTILQAFLESPCPPDITLSLKPTVS	479
Dd	419	EGRGERKKNDAPFAVSSGKRICVGBEALARMELFLYFTSILQRFSLRSUVPADIDIAHKIS	478
QY	480	GLFNPIPPAFLQ 490	
Dd	479	GFGNIPTPYEL 489	
RESULT 10			
CPB6 HUMAN			
ID	_CPB6 HUMAN	STANDARD;	PRT; 491 AA.
AC	P208I3;	Q9UK46;	
DT	01-FEB-1991	(Rel. 17, Created)	
DT	01-FEB-1991	(Rel. 17, Last sequence update)	
DT	10-OCT-2003	(Rel. 42, Last annotation update)	
DT	Cytochrome P450 2B6	(EC 1.14.14.1) (CYPIB6) (P450 IIB1).	
GN	CYP2B6.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OX	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OC	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RC	MEDLINE=90057429; PubMed=2573390;		
RX	Yamano S., Nambuuro P.T., Aoyama T., Meyer U.A., Inaba T., Kalow W.,		
PA	Gelboin H.V., McBride O.W., Gonzalez F.J.;		
RT	"cDNA cloning and sequence and cDNA-directed expression of human P450		
RT	IIB1: identification of a normal and two variant cDNAs derived from		
RT	the CYP2B locus on chromosome 19 and differential expression of the		
RT	IIB mRNAs in human liver."		
RL	Biochemistry 28:7340-7348(1989).		
RN	[2]		
RP	SEQUENCE FROM N.A., AND VARIANTS ALA-167; HIS-172 AND ARG-262.		
RC	TISSUE=Liver;		
RC	Zhuge J., Qian Y., Xie H., Yu Y.;		
RT	"Sequence of a new human cytochrome P450-2B6 cDNA."		
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
PA	Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,		
RA	Burkhardt-Schulz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,		
RA	Phan H., Velasco N., Do D., Regala W., Terry A., Brower A., Ganes J.,		
RA	Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,		
RA	Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefficient J.,		
RA	Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,		
RA	Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,		
RA	Olsen A.S., Carrano A.V.;		
RT	"Sequence analysis of a 4-Mb region in 19ql3.2."		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	TISSUE SPECIFICITY.		
RX	MEDLINE=20228924; PubMed=10768437;		
PA	Thum T., Borlak J.;		
RT	"Gene expression in distinct regions of the heart."		

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_taxID=10090;
 RN [1]
 RN TISSUE=Kidney;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA MEDLINE=89197954; PubMed=2703500;
 RA Lindberg R., Burkhardt B., Ichikawa T., Negishi M.;
 RT "The structure and characterization of type I P-450(15) alpha gene as
 RT major steroid 15 alpha-hydroxylase and its comparison with type II P-
 RT 450(15) alpha gene.";
 RL J. Biol. Chem. 264:6465-6471 (1989).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=17NC/Z;
 RA Jounaidi Y.;
 RT "cDNA and amino acid sequence of a new cyp2a isoform overexpressed in
 RT chemically induced mouse hepatoma.";
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN MUTAGENESIS.
 RP MEDLINE=89281737; PubMed=2733794;
 RX Lindberg R., Negishi M.;
 RA "Alteration of mouse cytochrome P450cch substrate specificity by
 RT mutation of a single amino-acid residue.";
 RL Nature 339:632-634 (1989).
 RN [4]
 RN TISSUE SPECIFICITY.
 RP MEDLINE=99421934; PubMed=1040589;
 RX Lavery D.J., Lopez-Molina L., Margueron R., Fleury-Olela F.,
 RA Conquet F., Schibler U., Bonfils C.;
 RT "Circadian expression of the steroid 15 alpha-hydroxylase (Cyp2a4) and
 RT coumarin 7-hydroxylase (Cyp2a5) genes in mouse liver is regulated by
 RT the PAR leucine zipper transcription factor DBP.";
 RL Mol. Cell. Biol. 19:6488-6499 (1999).
 CC -!- FUNCTION: EXHIBITS A HIGH COUMARIN 7-HYDROXYLASE ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- TISSUE SPECIFICITY: LIVER, WITH A STRONG CIRCADIAN RHYTHMICITY.
 CC -!- CIRCADIAN EXPRESSION IS REGULATED BY DBP.
 CC -!- DEVELOPMENTAL STAGE: IN LIVER; ACTIVITY 6 FOLD HIGHER IN FEMALES
 CC THAN IN MALES.
 CC -!- MISCELLANEOUS: There are only 11 differences between the sequence
 CC of testosterone 15-alpha-hydroxylase and that of coumarin 7-
 CC hydroxylase. By site-directed mutagenesis it has been shown that
 CC modification of position 209 is sufficient to convert the
 CC specificity of the two forms of the enzyme.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M25211; AAA37798.1; .
 CC EMBL; M26204; AAA37798.1; JOINED.
 CC EMBL; M25205; AAA37798.1; JOINED.
 CC EMBL; M25206; AAA37798.1; JOINED.
 CC EMBL; M25207; AAA37798.1; JOINED.
 CC EMBL; M25208; AAA37798.1; JOINED.
 CC EMBL; M25209; AAA37798.1; JOINED.
 CC EMBL; M25210; AAA37798.1; JOINED.
 CC EMBL; X89864; CAA61963.1; .
 CC PIR; B33531; B33531.
 CC HSSP; P00179; 1DT6.
 CC MGD; MGI:88597; Cyp2a5.
 CC InterPro; IPR001128; Cytochrome P450.
 CC InterPro; IPR008067; EP450_Cyp2A.

DR Pfam: PF00067; P450; 1.
 DR PRINTS; PR01684; EP450ICYP2A.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT METAL 439 439 IRON (HEME AXIAL LIGAND).
 SQ SEQUENCE 494 AA; 56740 MW; 1C2516D5FA2551D0 CRC64;
 Query Match 47.3%; Score 1237; DB 1; Length 494;
 Best Local Similarity 48.9%; Pred. No. 4.7e-79;
 Matches 239; Conservative 94; Mismatches 150; Indels 6; Gaps 3;
 QY 6 TWALLALALLLLLTALSGT----RARGHLPPTPLPLGLNLLQLRPGALYSGIMRLS 61
 DB 3 TSGLLLVAAVFLSVLMSVWKQKLSGKLPPTPLFFIGNFQLNTEQMYNSLMKIS 62
 QY 62 KYQGVFTIYIGMPVVLVQGEAVREALGQQAEEFSGRGIVAMLEGTFDGHGVFFSNG 121
 DB 63 QRYGVFTIYILGP-RRIVVLQGEAVKALVDQAEEFSGRGQATFDMLFKYGIVVFFSNG 121
 QY 122 EWRQLRKFTMLALRDGLGKREGELIQAEARCLVETPOGTGPRFPDPSLLLAQATSNV 181
 DB 122 ERAKQLRRPSTIATLDRDFGVGRKGIERIQEBAAGLIDSPKRTNGAFIDFTFVLSRVSNV 181
 QY 182 VCSLLFLGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHV 241
 DB 182 ISSIVFGDRFDYEDKEFLSLRLMMLGSGFTATSMGLQYEMFSVVMKHLPGPQQQAFKEL 241
 QY 242 STLAAPTVRQVOHQGNLDASGAPARDLVDAFLKKAQEQNPCTETNKNMLMTVLVLLF 301
 DB 242 QGLEDFITKKVBNHNTLDPNSP-RDFIDSFILMLEKKNPNTFYMKNLVLTLLNLF 300
 QY 302 AGTMVSTTVGYTLTLLMKYPHVQKWRELAELAGAGQAPSLGDRTRLPYTDVLAHEAQ 361
 DB 301 AGTETVSTTLRYGFLLMKHPDIEAKVHEEIDRVIGRNQPKYEDRMKMPYEAIVHEIQ 360
 QY 362 RLALLVPMGIPRTLMRTTRFRGYTLFPGTVEFPLLSILHDPNIFKHPBEFNPDRFLDAD 421
 DB 361 RADMLPMPGLARRVTKDFRDLFLPKGTVEFPLMSGLVKDKPFPSPKDFNPKHFLDDK 420
 QY 422 GPRKRHEAPLPSLGRVCLGEGAKAELFLFTTTILOAFSLESPPCPPTLSLKPVTSG 481
 DB 421 GQPKKDAFVPFSIGKRYCFEGELARMEFLFLTNIMQNPHFKSTQAPQDIDVSPRLVGF 480
 QY 482 FNIPPAFQL 490
 DB 481 ATIPPTITM 489
 RESULT 12
 CPBJ_MOUSE
 ID CPBJ_MOUSE STANDARD; PRT; 492 AA.
 AC OS5071;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 2B19 (EC 1.14.14.1) (CYP11B19).
 GN CYP2B19.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Skin;
 RX MEDLINE=99017986; PubMed=9799616;
 RA Keeney D.S.;
 RT "The novel skin-specific cytochrome P450 Cyp2b19 maps to proximal
 RT chromosome 7 in the mouse, near a cluster of Cyp2 family genes.";
 RL Genomics 53:417-419 (1998).
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an


```
CC -----
DR EMBL; M26208; AAA37797.1; -.
DR EMBL; M25146; AAA37797.1; JOINED.
DR EMBL; M25147; AAA37797.1; JOINED.
DR EMBL; M26202; AAA37797.1; JOINED.
DR EMBL; M26203; AAA37797.1; JOINED.
DR EMBL; M26205; AAA37797.1; JOINED.
DR EMBL; M26206; AAA37797.1; JOINED.
DR EMBL; M26207; AAA37797.1; JOINED.
DR EMBL; J03549; AAA40426.1; -.
DR EMBL; M19319; AAA40429.1; -.
DR PIR; A35331; A35331.
DR HSSP; P00179; 1DT6.
DR MGD; MGI:88596; Cyp2a4.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008067; EP450_CYP2A.
DR Pfam; PF00067; P450_1.
DR PRINTS; PR01684; EP450ICYP2A.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 439 439 IRON (HEME AXIAL LIGAND).
FT CONFLICT 236 236 L -> Q (IN REF. 2).
FT CONFLICT 306 306 V -> G (IN REF. 2).
SQ SEQUENCE 494 AA; 56594 MW; 69724BDE3195D75C CRC64;
Query Match 47.2%; Score 1235; DB 1; Length 494;
Best Local Similarity 49.1%; Pred. No. 6.4e-79;
Matches 240; Conservative 91; Mismatches 152; Indels 6; Gaps 3;
QY 6 TWALLALALLLLLTALSGT----RARGHLPPGPTPLPLGNLLQLRPGALYSGLMRLS 61
DB 3 TSGLLVAFAVFLSVLMSVWKQKLSGLKPLPGPTPLFVGNLQNTQMYNSLMKTS 62
QY 62 KYGVPFTYILGWPVPVVLVGOEAVREALGQAEFFSGRGVTVAMLEGTFDGHGVFSSG 121
DB 63 QYGGPFTYILGS--RRIVVLCGOEAVKEALVDQAEFFSGRGQATFDWLFKGYGIAFSSG 121
QY 122 ERWRQLRKFTMLALRDLGMKREGBELIQAEARCLVETTFQGTGRPFDPFSLLAQATSNV 181
DB 122 ERAKQLRSFSIATLDFGVGKRGIEBRIQEEAGFLIDSFRTKNGAFIDPTFYLSTVSNV 181
QY 182 VCSLLPGLRFSYEDKGFQAVRAAGTLLGVSSQGGQVTEMFSWFLRPLPGPHKQLLHV 241
DB 182 IGSIVGDFRFDYEDKFEFLSLLRMMLGSLQFTATSMGQVYEMFSVMKHLPGPQQQAFKEL 241
QY 242 STLAFTVRQVOHOGNLDASGPARDLVDFAFLKMAQEQNPGCTFTNKMMLTVIYLLF 301
DB 242 QGLEDFITRKVEHNRQTLDPNSP-RDFIDSLFRLMEKKNPTEFYMKXVLVLTTLNLF 300
QY 302 AGTMTVSTTVGYTLLLMKYPHYQVKVREBELNRELGAGQAPSLGDRTRLDYDAVLHEAQ 361
DB 301 AGTETVSTLYRGFLLLMKYPDIEAKVHEBIDRVGRNQPKYEDRDMKMPYTEAVIHEIQ 360
QY 362 RLIALVPMGIPRLMTRTRGCTYLPQGTVEFPLGSLILDHNPINFKHPESFNDRFLDAD 421
DB 361 RFADLIPMGHARRVTKDTRFDLLPKGTVEFVPMGLSVLKDPPKFFSNPKDFNFKHLDDK 420
QY 422 GRFKHEAFPLSFGKRVCLGEGLAKEALFFFTTILQAFSLSPSPDPTLSLKPTVSGL 481
DB 421 GQPKKSDAFVFFSIGKRYCFGEGLARMLFLFLTNMQNEHFKSTQAPQIDVSPRLVGF 480
QY 482 FNIPPAFQL 490
DB 481 VTIPPTVTM 489
RESULT 14
CPA3 RAT STANDARD; PRT; 494 AA.
ID CPA3 RAT
AC P20812;
DT 01-FEB-1991 (Rel. 17, Created)
```


Mon Sep 20 09:09:28 2004

Qy	357	LHEAQRLLALVPMGIPRTLMRTTRFRGYTLPQGTVEFPLGSLILHDPNIFKHPEEFNPDR	416
Db	356	IHEIQRTGDMPLMGLAHRVNDTKFRDFFLPKGTVEFPLGSELDRPRFFSNQDCSPQH	415
Qy	417	FLDADGRKHEAFPLPSLGRVCLGGLAKAELFLFFTILQAFSLSPCPCPDTLSLKP	476
Db	416	FLDEKGQFKSDAFVPPSIGKRYCFGEGLARWELFLFFTIMQNERFKSPQSPKXIDVSP	475
Qy	477	TVSGLFNIPPAFOLQVRP	494
Db	476	KHVGAFATIPRNYTWSFLP	493

Search completed: September 15, 2004, 09:19:00
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:14:41 ; Search time 43 seconds
(without alignments)
1127.454 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615
Sequence: 1 MEATGTWALLALALLLLT.....PPAFOLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2615	100.0	504	2 JC7613	cytochrome P450 2S
2	1319.5	50.5	494	2 B31944	cytochrome P450 2B
3	1313	50.2	491	2 S35666	cytochrome P450 2G
4	1300	49.7	491	1 O4RBP4	cytochrome P450 2B
5	1299	49.7	491	2 S31277	cytochrome P450 2B
6	1298	49.6	491	2 A27717	cytochrome P450 2B
7	1296	49.6	491	2 S31278	cytochrome P450 2B
8	1293	49.4	491	1 O4RTP2	cytochrome P450 2B
9	1289	49.3	491	1 O4RTPB	cytochrome P450 2B
10	1280.5	49.0	500	2 B31047	testosterone 16alp
11	1275	48.8	491	2 JT0676	cytochrome P450 2B
12	1262	48.3	494	2 S11305	cytochrome P450 2B
13	1247	47.7	491	2 I84735	testosterone 16alp
14	1244	47.6	491	2 A31047	cytochrome P450 2B
15	1243.5	47.6	494	2 A35551	testosterone 16a-h
16	1238	47.3	491	2 A32969	cytochrome P450 2G
17	1237	47.3	494	2 B33531	cytochrome P450 2B
18	1235	47.2	494	2 A33531	cytochrome P450 2A
19	1234.5	47.2	494	2 A32030	testosterone 15alp
20	1220	46.7	494	2 S16068	coumarin 7-monooxy
21	1213.5	46.4	494	1 O4HUA6	testosterone 16a-h
22	1210.5	46.3	492	2 S27160	cytochrome P450 2A
23	1210	46.3	491	2 I49625	cytochrome P450 2A
24	1203.5	46.0	494	2 A47494	cytochrome P450 2A
25	1198.5	45.8	494	2 B47494	cytochrome P450 2A
26	1175.5	45.0	494	2 I38965	cytochrome P450 2A
27	1174.5	44.9	494	2 C34271	cytochrome P450 50
28	1168	44.7	491	2 A39302	cytochrome P450 ar
29	1163.5	44.5	494	2 A46588	cytochrome P450 ar

30 1160 44.4 491 2 A29818 cytochrome P450 2B
31 1153.5 44.1 494 2 I38967 cytochrome P450 -
32 1150.5 44.0 492 2 S32491 testosterone 7alpb
33 1144.5 43.8 492 2 A34272 testosterone 7alpb
34 1108 42.4 491 2 A36036 cytochrome P450 2F
35 1107 42.3 491 2 A31418 cytochrome P450 2H
36 1105.5 42.3 494 2 A33293 cytochrome P450 2A
37 1098 42.0 487 1 OARBC6 cytochrome P450 2C
38 1097.5 42.0 490 2 B28516 progesterone monoo
39 1096 41.9 487 1 OARBP4 testosterone 7alpb
40 1091.5 41.7 492 2 A31887 testosterone 7alpb
41 1090.5 41.7 494 2 S31301 S-mephenytoin 4-hy
42 1085.5 41.5 490 2 B38462 cytochrome P450 CY
43 1083 41.4 491 2 A24814 cytochrome P450 -
44 1079.5 41.3 490 2 I48163 cytochrome P450 -
45 1078.5 41.2 490 2 D28951 cytochrome P450 2C

ALIGNMENTS

RESULT 1

JC7613

cytochrome P450 2S1 protein, CYP2S1 - human

C:Species: Homo sapiens (man)

C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Jun-2002

C:Accession: JC7613

R:Rylander, T.; Neve, E.P.A.; Ingelman-Sundberg, M.; Oscarson, M.

Biochem. Biophys. Res. Commun. 281, 529-535, 2001

A:Title: Identification and tissue distribution of the novel human cytochrome P450 2S1 ((

A:Reference number: JC7613; MUID:21092856; PMID:11181079

A:Contents: Liver

A:Accession: JC7613

A:Molecule type: mRNA

A:Residues: 1-504 <RYL>

A:Cross-references: GB:AF335278

C:Comment: This protein, a novel member of cytochrome P450 2(CYP2) family, has a role in

C:Genetics:

A:Gene: cyp2s1

A:Map position: 19q13.2

A:Introns: 59/3; 115/1; 165/1; 218/3; 278/3; 326/1; 388/3; 436/1

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C:Keywords: heme; iron; liver; metalloprotein

F:33-36/Region: conserved PPGP sequence #status predicted

F:440/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 100.0%; Score 2615; DB 2; Length 504;

Best Local Similarity 100.0%; Pred. No. 3.5e-205;

Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTALSCTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSCTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60
QY 61 SKKYGPFYIYLGPNRPVVVLVQGEAVREALGQAEFEFGSGRTVAMLEGTFDGHGVFFSN 120
Db 61 SKKYGPFYIYLGPNRPVVVLVQGEAVREALGQAEFEFGSGRTVAMLEGTFDGHGVFFSN 120
QY 121 GERWRQLRFTMLALRDLGMRKEGELTQAEARCLVETFOGTGPRPDPSPSLLAATS 180
Db 121 GERWRQLRFTMLALRDLGMRKEGELTQAEARCLVETFOGTGPRPDPSPSLLAATS 180
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSGGQGYEMFSLRPLRPGHKOLLLH 240
Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSGGQGYEMFSLRPLRPGHKOLLLH 240
QY 241 VSTLAFTVROVQOQHGNDASGPARDLVDAFLKMAQEEQNPGETFNKNMLMTVIYLL 300
Db 241 VSTLAFTVROVQOQHGNDASGPARDLVDAFLKMAQEEQNPGETFNKNMLMTVIYLL 300
QY 301 FAGTMTVSTTGYTTLILLMKYPHVQKVRRELNRELGAQPSLGDRTLPYTDVLAHEA 360
Db 301 FAGTMTVSTTGYTTLILLMKYPHVQKVRRELNRELGAQPSLGDRTLPYTDVLAHEA 360

Db 359 IQRLTDVPMGVPHVIRDTFRGYLLPKGTDPVLLGSLVKOPKFCPPDDFPQHFELD 418
 Qy 420 ADGRFKHEAFPLPSLGRVCLGELAKAEFLFETTLQAFSLSPCPDPTLSLKPTVS 479
 Db 419 EQGRFKNEAFVPPSSKRICLGEMARMELFLFTSLQNFSLHPLVPFNIDITPKIS 478
 Qy 480 GLFNIPPAFOL 490
 Db 479 GFGNIPPTIEL 489
 RESULT 3
 S35666
 Cytochrome P450 2B4 isoform Bx - rabbit
 N/Alternate names: cytochrome P450 2B isoform Bx
 N/Contains: oxidoreductase (EC 1.-.-.-)
 C/Species: Oryctolagus cuniculus (domestic rabbit)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 04-Mar-2000
 C/Accession: S35666; B54251
 R/Ryan, R.; Grimm, S.W.; Kedzie, K.M.; Halpert, J.R.; Philpot, R.M.
 Arch. Biochem. Biophys. 304, 454-463, 1993
 A/Title: Cloning, sequencing, and functional studies of phenobarbital-inducible forms of
 A/Reference number: S35666; MUID:93349061; PMID:8346920
 A/Accession: S35666
 A/Molecule type: mRNA
 A/Residues: 1-491 <RYA>
 A/Cross-references: EMBL:S64259; NID:G402842; PIDN:AAB27705.1; PID:G402843
 R/Roberts, E.S.; Hopkins, N.E.; Zaluzec, E.J.; Gage, D.A.; Alworth, W.L.; Hollenberg, P.
 Biochemistry 33, 3766-3771, 1994
 A/Title: Identification of active-site peptides from (3)H-labeled 2-ethynylaphthalene-1
 A/Reference number: A54251; MUID:94190899; PMID:8142377
 A/Accession: B54251
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 273-283 <ROB>
 C/Genetics:
 A/Gene: CYP2B4
 C/Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homolog
 C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
 F/295-458/Domain: cytochrome P450 homolog <P45>
 F/436/Binding site: heme iron (Cys) (axial ligand) #status predicted
 Query Match 50.2%; Score 1313; DB 2; Length 491;
 Best Local Similarity 50.7%; Pred. No. 4.6e-99;
 Matches 247; Conservative 99; Mismatches 137; Indels 4; Gaps 3;
 Qy 7 WALLALALL--LLLTALSGTRAGHLPPTPLPLGNLLQRLPGALYSGLMRLSKY 64
 Db 3 FSLLLALLAFLAGLLLLFRGHKPAHGRLLPPGPSPLVGLNLLQMDRKLRLRUREKY 62
 Qy 65 GPVFTYILGPWRPVVVLVQEAAREALGQAEFEFGRTVAMLEGTDPGHGVFSGNERW 124
 Db 63 GDVFTVYLSG--RPVVVLCGTDAIREALVDQEAFAFGKGIADVDPFQGGVIFANGERW 121
 Qy 125 RQLRKFTMLALDLGKMGKREGELIQAEARCLVETFOGTGPRPDPFSLLAQATSNVCS 184
 Db 122 RALRRFSLATRDPFGMGKRSVEERIQEAARCLVEELRKSGLALLDNTLLFHSITSNICS 181
 Qy 185 LLEGLRFSYEDKFEQAVVRAAGTLLGVSSQGGQTYEMFWSFLPLPGPHKQLLHHVSTL 244
 Db 182 IVFGKFDYKDPVFLRLDLLFQSFSLISSFSQVFLFSGFLKHFPGTHQIRYNIQEI 241
 Qy 245 AAFVROVQOQHNLDASGPARDLVDAFLKMAEQEONPGTEFTNKMLMTVIYLLFAGT 304
 Db 242 NTFITQSVKHAATLDPSNP--RDFIDVYLRMEKDKSDPSSEFHQNLILTVLSFPAGT 300
 Qy 305 MTVSTTVGYTLLLMKYPHVQKVRRELNRELGAQAPSLGDRTRLPYTDAVLHEAQRLL 364
 Db 301 ETTSTTLRYGLLMKYPHVQKVRRELNRELGAQAPSLGDRTRLPYTDAVLHEAQRLL 360
 Qy 365 ALVPMGIPRLMRTTRFRGYTLPGQTEVFFLLGSLILHDPNIFKHPEEFPDRFLD 424

Qy 361 QRLALVPMGIPRLMRTTRFRGYTLPGQTEVFFLLGSLILHDPNIFKHPEEFPDRFLD 420
 Db 361 QRLALVPMGIPRLMRTTRFRGYTLPGQTEVFFLLGSLILHDPNIFKHPEEFPDRFLD 420
 Qy 421 DGRFKHEAFPLPSLGRVCLGELAKAEFLFETTLQAFSLSPCPDPTLSLKPTVS 480
 Db 421 DGRFKHEAFPLPSLGRVCLGELAKAEFLFETTLQAFSLSPCPDPTLSLKPTVS 480
 Qy 481 LFNIPPAFOLVPTDLHSTTQTR 504
 Db 481 LFNIPPAFOLVPTDLHSTTQTR 504
 RESULT 2
 B31944
 Cytochrome P450 2G1 - rabbit
 N/Alternate names: cytochrome P450NMB
 N/Contains: oxidoreductase (EC 1.-.-.-)
 C/Species: Oryctolagus cuniculus (domestic rabbit)
 C/Date: 31-Mar-1990 #sequence_revision 17-May-1996 #text_change 04-Mar-2000
 C/Accession: S31907; B31944
 R/Ding, X.; Porter, T.D.; Peng, H.M.; Coon, M.J.
 Arch. Biochem. Biophys. 285, 120-125, 1991
 A/Title: cDNA and derived amino acid sequence of rabbit nasal cytochrome P450NMB (P450I1
 A/Reference number: S31907; MUID:91119408; PMID:1703755
 A/Accession: S31907
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 12-494 <DIN>
 R/Ding, X.; Coon, M.J.
 Biochemistry 27, 8330-8337, 1988
 A/Title: Purification and characterization of two unique forms of cytochrome P-450 from
 A/Reference number: A90541; MUID:89207441; PMID:3242590
 A/Accession: B31944
 A/Molecule type: protein
 A/Residues: 1-45, 'E', 47-50 <D12>
 C/Genetics:
 A/Gene: CYP2G1
 C/Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homolog
 C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
 F/298-461/Domain: cytochrome P450 homolog <P45>
 F/439/Binding site: heme iron (Cys) (axial ligand) #status predicted
 Query Match 50.5%; Score 1319.5; DB 2; Length 494;
 Best Local Similarity 50.5%; Pred. No. 1.4e-99;
 Matches 248; Conservative 97; Mismatches 143; Indels 3; Gaps 3;
 Qy 1 MEATGATWALLALAL--LLLTALSGTRAGHLPPTPLPLGNLLQRLPGALYSGLMR 59
 Db 1 MELGGAFTIFLALCFSCLLILIAWKRVQKFGRLPPGPTPIFFGNLLQVTDATFQSLK 60
 Qy 60 LSKYGPVFTYILGPWRPVVVLVQEAAREALGQAEFEFGRTVAMLEGTDPGHGVFSS 119
 Db 61 LREKYGVTYVYMP--RPVVILCHEAVKALVDRADEFSGRGLASVERNFQGHGVALA 119
 Qy 120 NGERWRQLRFTMLALDLGKMGKREGELIQAEARCLVETFOGTGPRPDPFSLLAQATS 179
 Db 120 NGERWRLRPSLTILADFGMGKSEIERIQEAAGYLLFEFRKTKGAPIDPTFFLSRTVS 179
 Qy 180 NVVCSLLFGRLFSYEDKFEQAVVRAAGTLLGVSSQGGQTYEMFWSFLPLPGPHKQLLH 239
 Db 180 NVISSVVFGRFYEDKQFSLRLMINESFIEMSTPWAQLYDMYGVQVYLPFGHRIYI 239
 Qy 240 HVSTLAFTVROVQOQHNLDASGPARDLVDAFLKMAEQEONPGTEFTNKMLMTVIYL 299
 Db 240 LIEELKDFIARVKNVNEASLDPNP--RDFIDCFILKHKQDKNPHTEFNKLVLTLLN 298
 Qy 300 LFACTMTVSTTVGYTLLLMKYPHVQKVRRELNRELGAQAPSLGDRTRLPYTDAVLHE 359
 Db 299 PFAGTETVSSLTRYGFLIMKHPEVQTKIYEINQVGPRIIPSVDDVRKMPPTDAVHE 358
 Qy 360 AQRLLALVPMGIPRLMRTTRFRGYTLPGQTEVFFLLGSLILHDPNIFKHPEEFPDRFLD 419

Db 361 DIPIFGVHTVTKDQFRGVIPKNTVEPVVSSALHDPYKTPNTFNGHFDANGAL 420
QY 425 RKHEAFPSLCKRYCLGSLAKAELEFPTTILQAFSLSPDPTLSLKTPTVSLFNI 484
Db 421 KNEGFMPFSLCKRYCLGSLAKAELEFPTTILQAFSLSPDPTLSLKTPTVSLFNI 484
QY 485 PPAFQLQ 491
Db 481 PPSYQIR 487

RESULT 4
Cytochrome P450 2B4 - rabbit
N:Alternate names: cytochrome P450-1M2
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 20-Sep-1984 #sequence revision 20-Sep-1984 #text change 03-Mar-2000
C:Accession: A00179; A61538; S31279; A00178; B27717; C27717; E27717
R:Tarr, G.E.; Black, S.D.; Fujita, V.S.; Coon, M.J.
Proc. Natl. Acad. Sci. U.S.A. 80, 6552-6556, 1983
A:Title: Complete amino acid sequence and predicted membrane topology of phenobarbital-1
A:Reference number: A00179; MUID:84042509; PMID:6579541
A:Accession: A00179
A:Molecule type: protein
A:Residues: 1-491 <TAR>
R:Parandooch, Z.; Fujita, V.S.; Coon, M.J.; Philpot, R.M.
Drug Metab. Dispos. 15, 59-67, 1987
A:Title: Cytochrome P-450 isozymes 2 and 5 in rabbit lung and liver. Comparisons of str
A:Reference number: A61538; MUID:87161284; PMID:2881760
A:Accession: A61538
A:Molecule type: protein
A:Residues: 1-24 <PAR>
R:Gasser, R.; Negishi, M.; Philpot, R.M.
Mol. Pharmacol. 33[32], 22-30, 1988
A:Title: Primary structures of multiple forms of cytochrome P-450 isozyme 2 derived from
A:Reference number: S31277
A:Note: header on page 22 gives volume number as 32
A:Accession: S31279
A:Molecule type: mRNA
A:Residues: 1-491 <GAS>
A:Cross-references: EMBL:M20856; NID:g164958; PIDN:AAA65840.1; PID:g164959
R:Heinemann, F.S.; Ozols, J.
J. Biol. Chem. 258, 4195-4201, 1983
A:Title: The complete amino acid sequence of rabbit phenobarbital-induced liver microsom
A:Reference number: A00178; MUID:83160983; PMID:6833251
A:Accession: A00178
A:Molecule type: protein
A:Residues: 1-90, 'E', 92-94, 96, 95, 97-98, 101-134, 'GY', 137-140, 'G', 142-192, 'K', 194-220, 'S',
R:Komori, M.; Imai, Y.; Tsunashima, S.; Sato, R.
Biochemistry 27, 73-80, 1988
A:Title: Microheterogeneity in the major phenobarbital-inducible forms of rabbit liver m
A:Reference number: A90538; MUID:88163620; PMID:2831964
A:Accession: B27717
A:Molecule type: mRNA
A:Residues: 301-313, 'L', 315-419, 'M', 421-491 <KOM>
A:Experimental source: clone b14
A:Accession: C27717
A:Molecule type: mRNA
A:Residues: 52-103, 'M', 105-173, 'V', 175-220, 'S', 222-313, 'L', 315-424, 'C', 426-491 <KO2>
A:Experimental source: clone b46
A:Note: the authors translated the codon AAG for residue 191 as Arg
A:Accession: E27717
A:Molecule type: mRNA
A:Residues: 297-479, 'L', 481-491 <KO3>
A:Experimental source: clone b54
C:Comment: Cytochromes P450 are a group of membrane-bound hemoprotein monooxygenases. In
variety of structurally unrelated compounds, including steroids, fatty acids, and xenob
C:Genetics:
A:Gene: CYP2B4
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: cytochrome; electron transfer; endoplasmic reticulum; heme; iron; metallo

F:295-458/Domain: cytochrome P450 homology <P45>
F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 49.7%; Score 1300; DB 1; Length 491;
Best Local Similarity 50.1%; Pred. No. 5.3e-98;
Matches 244; Conservative 101; Mismatches 138; Indels 4; Gaps 3;

QY 7 WALLALALL--LLLTALSGTRARGHLPFGTPTPLGLNLLQLRPGALYSGLMRLSKY 64
Db 3 FSLLLLAFLAGLLLLFRGHPKAGHRLPPGSPPLVGLNLLQMDRKGLLRFLRLREKY 62
QY 65 GPVFTIYLGPRPVVYVNGQEAVERALGQAEFSGRGTVMLEGTGFGHVFNSGERM 124
Db 63 GDVFTVILGS-RPVVVLGCTDAIREALVDQAEAFSGRGTIAVVDPIFGYGVIFANGERW 121
QY 125 RQLRKFTMLALRDLGMKREGEELIQAARCLIVETFGTEGTPFPDPSLLLAQATSNVVC 184
Db 122 RALRRFSLATWRDFGMRKRSVEERIQEAEARCLIVELRKSAGALLDNTLLFHSITSNICS 181
QY 185 LLFGLRFSYEDKEFOAVVRAAGGTLILGVSSQGGQTYEMFSLRPLPGPHKOLLHHVSTL 244
Db 182 IVFGKRFYKDPVFLRLDLDFQSFSLISSFSQVFLFPGFLKHFPGTHRQIYRNLOEI 241
QY 245 AAFVTRVQOQGNLDASGPARDLVDAFLKMAQEEQNPTEFTNKNMLMTVIYLLFAGT 304
Db 242 NTFIGQSVEKHATLDPSNP-RDFIDVYLLRMEKDKSPSEFHHQNLILATVLSFFAGT 300
QY 305 MTVSTTVGYTLLLMKYPHVOKWVREELNRELGAQAPSLGDRTPVTDVAVLHEAQLL 364
Db 301 ETTTILRYGFLMLKYPHVTERVQKIEQVIGSHRPPALDDRAKMPYTDVAVIHIQRLG 360
QY 365 ALVPMGIPRLMTRFRGVTLPGTEVFPVLLGSLILHDPNIFKHPEFNPORFLDADGRF 424
Db 361 DLIPFGVHTVTKDQFRGVIPKNTVEPVVSSALHDPYKTPNTFNGHFDANGAL 420
QY 425 RKHEAFPSLCKRYCLGSLAKAELEFPTTILQAFSLSPDPTLSLKTPTVSLFNI 484
Db 421 KNEGFMPFSLCKRYCLGSLAKAELEFPTTILQAFSLSPDPTLSLKTPTVSLFNI 484
QY 485 PPAFQLQ 491
Db 481 PPSYQIR 487

RESULT 5
S31277
cytochrome P450 2B4-B1 - rabbit
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 28-May-1993 #sequence revision 28-May-1993 #text_change 04-Mar-2000
C:Accession: S31277; I46699
R:Gasser, R.; Negishi, M.; Philpot, R.M.
Mol. Pharmacol. 33[32], 22-30, 1988
A:Title: Primary structures of multiple forms of cytochrome P-450 isozyme 2 derived from
A:Reference number: S31277
A:Accession: S31277
A:Molecule type: mRNA
A:Residues: 1-491 <GAS>
A:Cross-references: EMBL:M20857; NID:g164960; PIDN:AAA31224.1; PID:g164961
A:Gene: CYP2B4
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:295-458/Domain: cytochrome P450 homology <P45>
F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 49.7%; Score 1299; DB 2; Length 491;
Best Local Similarity 50.1%; Pred. No. 6.4e-98;
Matches 244; Conservative 101; Mismatches 138; Indels 4; Gaps 3;

QY 7 WALLALALL--LLLTALSGTRARGHLPFGTPTPLGLNLLQLRPGALYSGLMRLSKY 64
Db 3 FSLLLLAFLAGLLLLFRGHPKAGHRLPPGSPPLVGLNLLQMDRKGLLRFLRLREKY 62

QY	64	YGVFTIYIGPWRPVVVLVGQEAUREALGQAAEEFSGRGRTVAMLEGTFDCHGVFFSNGR	123
Db	62	YGVFTIYVLGS-RPVVVLCTDAIREALVDQAEAFSGRGKIAVLDPFQGVYTFANGER	120
QY	124	WROIRFTMLALBDLGMGKEEGEELIQAEARCLIVETFOGTGEGRPFPDPSLILQAQATSNVVC	183
Db	121	WRALRRFSLATMRDFFGMGKKSVEERIQEARCLIVEELRKSAGALLDNTLLFHSVTSNIC	180
QY	184	SLIPLGLRFSYEDKEFOAVVRAAGGTLIGVSSQGGQYVEMFSWFLRPLRPGHKQLLHVST	243
Db	181	SIVFGKEFDYKDPVFLKLLDLFFQSPSLTSSFSQVPELFSGLFKHPGTHROIYRNLOE	240
QY	244	LAAPTVOQOCHQNLDSAGPARDLVDAFLIKWAAEQBNPCTBTKNMLMTVILYLLPAG	303
Db	241	INTIGOSVHRHRTDPSNP-RDFIDVILRLMEKOKSDPSSEPHHQNLILTVLTLFPAG	299
QY	304	TMTVSTTVGYTLILLIMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDVLUHEAORL	363
Db	300	TEITSTTLRYGFLILLKYPHWTERVQKEIEQVIGSHRPPALDDRAKMPYTDVATHEIQR	359
QY	364	LALVPMGIDPRLMTTTRFRGYTLPOGTEVPLGSLIILDANI FKHPBEFNPDRELDADR	423
Db	360	GDLIPLGYPHWMTXDTQFRGVIPKNTIEVPVLSSALHDPRYFETPTNFPNGHFLDADGA	419
QY	424	FRKEAFUPLFSLGKRVCLGEAKAEFLFPTTTLQAFSLIESPCPPDTLSUKPTVSLGFN	483
Db	420	LKRIEGFWPFSLGKRI CLGEGIAETLFLPTTTLQONFSIASPVPEPIDILTPRESGVGN	479
QY	484	IPPAFQLQ 491	
Db	480	VPPSYQIR 487	

[illegible]

Db 63 GDFVTVYLGSRPVPVVLGCTDAIREALVDOAEAFSGRGIKIAVVDPIFGQYGVFFANGEHM 121
 QY 125 ROLKFTMLARDLGMKREGEELLOAARCLIVETFGQTEGRPPDPSLLLAQATSNVCS 184
 Db 122 RALRRFSLSATWRDFGMGRSVEERIOEAEARCLIVEELRSKGALLDNTLLFHSVTSNIICS 181
 QY 185 LLFGLRFSYEDKEFAVVRAAGGTLGVSSGGQTYEMFWSFLRLPLCPGPHKOLLHHVSTL 244
 Db 182 IVFGKRPYKDPVFLRLDLDFQSFSLISSRSSQVFLFPGLFKHFGPTHQIQRVNLQEI 241
 QY 245 AAFVTRVQVQOQGNLDGSGPARDLVDAFLKVAQEQNPQTEFTNKNMLMTVILLFAGT 304
 Db 242 NTFITGQVTEKRRATLDSPN-RDFIDVLLRMKXKSDPSSEFHRNLLIVTLVTLFFAGT 300
 QY 305 MIVSTTVGYTLLMLKYPHVQKWRLEENRLGAGQAPSLGDRTRLPTDVAHLHQAQLL 364
 Db 301 ETTSTTURYGFLLMLKYPHVQKWRLEENRLGAGQAPSLGDRTRLPTDVAHLHQAQLL 360
 QY 365 ALVPMGIPTLMTRFRGYTLPGQTEVFFLLGSLHDPNIFKHPENPDRELDADGRF 424
 Db 361 DLVPFGAPHMVKDTQFRGVIPKNTVEVPLVSSALHDPFRYFETNTFNPQGHFLDAGAL 420
 QY 425 RKHEAPFSLGKRVCLGEGAKAELELFFTTILOAFSLSPCPDPTLSLKPVTSGLFNI 484
 Db 421 KRNEGFMFSLGKRVCLGEGAKAELELFFTTILOAFSLSPCPDPTLSLKPVTSGLFNI 480
 QY 485 PPAFLQLQ 491
 Db 481 PPSYQIR 487

RESULT 8
 O4RTP2
 Cytochrome P450 2B2 - rat
 N/Alternate names: cytochrome P450 PB-4; cytochrome P450, phenobarbital-inducible; cytochrome P450 2B2
 C/Species: Rattus norvegicus (EC 1.1.1.-)
 C/Date: 04-Dec-1986 #sequence revision 17-May-1996 #text change 01-Dec-2000
 R/Mizukami, Y.; Sogawa, K.; Suwa, Y.; Muramatsu, M.; Fujii-Kuriyama, Y.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3958-3962, 1983
 A/Title: Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver.
 A/Reference number: A21162; MUID:83247397; PMID:6306654
 A/Accession: A21162
 A/Molecule type: DNA
 A/Residues: 1-472, 'M', 474-491 <MIT>
 A/Cross-references: EMBL:J00728; NID:g203845; PIDN:AAA1056.1; PID:g203847
 A/Note: the authors translated the codon AGT for residue 4 as Thr, and ATG for residue 3 as Thr.
 R/Frey, A.B.; Waxman, D.J.; Kreibich, G.
 J. Biol. Chem. 260, 15253-15265, 1985
 A/Title: The structure of phenobarbital-inducible rat liver cytochrome P-450 isoenzyme H
 A/Reference number: A00177; MUID:86059379; PMID:3877725
 A/Accession: A00177
 A/Molecule type: protein
 A/Residues: 1-291, 'P', 293-320, 'AE', 323-475, 'D', 477-491 <PRE>
 R/Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
 Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982
 A/Title: Primary structure of a cytochrome p-450: coding nucleotide sequence of phenobarbital-inducible rat liver cytochrome P-450
 A/Reference number: A93912; MUID:82222224; PMID:6953431
 A/Accession: B00176
 A/Molecule type: mRNA
 A/Residues: 6-359, 'S', 361-362, 'V', 364-366, 'V', 368-406, 'S', 408-416, 'N', 418, 'A', 420-477, 'G'
 A/Note: nucleotide sequence for residues 1-5 is not given
 A/Note: the authors translated the codon GAT for residue 166 as Glu, CTG for residue 292 as Glu, and the authors translated the codon GAT for residue 166 as Glu, CTG for residue 292 as Glu.
 R/Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
 Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982
 A/Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenobarbital-inducible rat liver cytochrome P450
 A/Reference number: A93925
 A/Cross-references: EMBL:J00728; NID:g203845; PIDN:AAA1056.1; PID:g203847
 A/Note: the mistranslations in reference A93912 are acknowledged
 R/Botelho, L.H.; Ryan, D.B.; Levin, W.
 J. Biol. Chem. 254, 5635-5640, 1979

A/Title: Amino acid compositions and partial amino acid sequences of three highly purified or 3-methylcholanthrene.
 A/Reference number: A92255; MUID:79194111; PMID:109438
 A/Accession: B92255
 A/Molecule type: protein
 A/Residues: 1-3, 'T', 5-22 <BOT>
 R/Lacroix, D.; Desrochers, M.; Lambert, M.; Anderson, A.
 Gene 86, 201-207, 1990
 A/Title: Alternative splicing of mRNA encoding rat liver cytochrome P450e (P450IIB2).
 A/Reference number: S15589; MUID:90215299; PMID:2323573
 A/Accession: S15589
 A/Molecule type: mRNA
 A/Residues: 105-113, 'F', 115-274, 'VSPAMRE', 275-321, 'E', 323-491 <IAC>
 A/Cross-references: EMBL:M34452; NID:g203679; PIDN:AAA1004.1; PID:g203680
 A/Note: translation of the nucleotide sequence is not complete
 R/Philips, I.R.; Shephard, E.A.; Ashworth, A.; Rabin, B.R.
 Gene 24, 41-52, 1983
 A/Reference number: A21872
 A/Accession: A21872
 A/Molecule type: mRNA
 A/Residues: 168-321, 'E', 323-443, 'K', 445-491 <PHI>
 R/Afolter, M.; Anderson, A.
 Biochem. Biophys. Res. Commun. 118, 655-662, 1984
 A/Title: Segmental homologies in the coding and 3' non-coding sequences of rat liver cytochrome P450
 A/Reference number: A32736; MUID:84153837; PMID:6322758
 A/Accession: A32736
 A/Molecule type: mRNA
 A/Residues: 385-491 <AFF>
 A/Cross-references: GB:K01626; NID:g203782; PIDN:AAA41037.1; PID:g203783
 R/Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.
 Arch. Biochem. Biophys. 270, 23-32, 1989
 A/Title: Antibodies targeted against hypervariable and constant regions of cytochromes P450
 A/Reference number: S03854; MUID:89192373; PMID:2539047
 A/Accession: S03854
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 329-358, 'AS', 361, 362, 363-380, 402-423 <OES>
 R/Atchison, M.L.; Adesnik, M.
 Proc. Natl. Acad. Sci. U.S.A. 83, 2300-2304, 1986
 A/Title: Gene conversion in a cytochrome P-450 gene family.
 A/Reference number: I59060; MUID:86205943; PMID:3458196
 A/Accession: I59060
 A/Status: preliminary; translated from GB/EMBL/DDBU
 A/Molecule type: DNA
 A/Residues: 323-431 <RES>
 A/Cross-references: GB:M13234; NID:g203848; PIDN:AAA41057.1; PID:g554434
 C/Genetics:
 A/Gene: CYP2B2
 A/Introns: 384/3
 C/Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
 C/Keywords: alternative splicing; chromoprotein; electron transfer; endoplasmic reticulum
 F/295-458/Domain: cytochrome p450 homology <P45>
 F/436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 49.4%; Score 1293; DB 1; Length 491;
 Best Local Similarity 50.4%; Pred. No. 2e-97;
 Matches 243; Conservative 96; Mismatches 141; Indels 2; Gaps 2;

QY 9 LLLALLLLLLTLLSGTFAHGLPPTPLGLNLLQLRPGALYSGMLRLSKYGPVF 68
 Db 7 LLLALLVGLLLLVGHGPKSRGNFPGRPLGLNLLQLDRGGLNSFMQLREKYGVDF 66

QY 69 TYLGPWRPVVVLGVQAEVREALGQAEFFSGRGTVAMLEGTFDGHGVFFSGNWRQLR 128
 Db 67 TVHLGF-PPVVMCGTDTIKALVQAEDFSGRGIVAVIEPIFKEYGVIFANGERKALR 125

QY 129 KFTMALRDLGMKREGEELIOAEARCLIVETFGQTEGRFPDPSLLLAQATSNVCSLLFG 188
 Db 126 RFSLATMRDFGMGRSVEERIOEAEARCLIVEELRSKGALLDNTLLFHSVTSNIICSIVFG 185

QY 189 LRFSYEDKEFAVVRAAGGTLGVSSGGQTYEMFWSFLRLPLCPGPHKOLLHHVSTLAAPT 248
 Db 186 ERFDYTDQFLLELFFRTLSLSSFSQVFEFFSGFLKYPFGAHRQIKNLQIBLDYI 245

RESULT 10

B31047
testosterone 16alpha-hydroxylase (EC 1.14.14.14.-) cytochrome P450 2B10 - mouse
N:Alternate names: cytochrome P450CDB; cytochrome P450pf3/46
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 04-Mar-2000
C:Accession: B31047; A60559
R:Noshiro, M.; Iakso, M.; Kawajiri, K.; Negishi, M.
Biochemistry 27, 6434-6443, 1988
A:Title: Rip locus: regulation of female-specific isozyme (I-P-450-16-alpha) of testosterone
A:Reference number: A31047; PMID:89118235; PMID:3219345
A:Accession: B31047
A:Molecule type: mRNA
A:Residues: 1-500 <NO>
A:Cross-references: EMBL:M21856; NID:g201968; PIDN:AAA0425.1; PID:g201969
A:Note: the authors translated the codon TTA for residue 55 as Phe, AGA for residue 133 as Val, AAA for residue 236 as Asn, AGG for residue 251 as Lys, AAG for residue 253 as 3 as Asp
A:Note: the authors translated the codon GCA for residue 281 as Thr, TTC for residue 283 as Arg, ACA for residue 321 as Ala, GTG for residue 331 as Leu, TCA for residue 349 as s Ala, and GAC for residue 479 as Gly
A:Note: the sequence nucleotide translation from Fig. 8 is inconsistent with the nucleot R:Bornheim, L.M.; Correia, M.A.
Mol. Pharmacol. 36, 377-383, 1989
A:Title: Purification and characterization of a mouse liver cytochrome P-450 induced by A:Reference number: A60559; PMID:89384475; PMID:2779523
A:Accession: A60559
A:Molecule type: protein
A:Residues: 1-15 <BO>
A:Note: this enzyme was induced by cannabidiol
C:Genetics:
A:Gene: Cyp2b-10
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo F:295-467/Domain: cytochrome P450 homology <P45>
F:445/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 49.0%; Score 1280.5; DB 2; Length 500;
Best Local Similarity 49.7%; Pred. No. 2.1e-96;
Matches 244; Conservative 99; Mismatches 137; Indels 11; Gaps 3;
QY 9 LLLALLALLLLTALSGTRAGHLPDPTPLPILGNLQLRPGALYSGLMRLSKYGFV 68
DB 7 LLLALLVGLLLLARGHPSKRGNGPPGPRPLPILGNLQMDRGLLKSLQLREKYGVF 66
QY 69 TIYLGPRPVVVLVGOEAVREALGGQAEESGRTVAMLEGTGCHGVFNSGGERWQLR 128
DB 67 TVHLGP-RPVVLLCGDITREALVGOEAVREALGGQAEESGRTVAMLEGTGCHGVFNSGGERWQLR 125
QY 129 KFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPDPSLLLAQAATSNVVCSLFG 188
DB 126 RFSLATMRDGMKRSVEERIQEAEQCLVEELRKSQAPLDPTFLQCITANVICSIVFG 195
QY 189 LRFSEDEKFEQAVVRAAGTLLGVSSQGGQTYEMFSLRPLRPGPKQLLHVSTLAAFT 248
DB 186 ERFETDRLPRLLELFYQTFSLISSFSQMFELFSGFLKYFPCAHRQISKNLQELLDYI 245
QY 249 VRQVOQHGNLDASGPARDLVDAFLKMAQEEONPGTEFTKNMMLTVIYLLFAGTMTVS 308
DB 246 GHSVERKATLDPSVP-RFIDILYLRMEKESQNAEFHQMLMMSVLSPFVGTST 304
QY 309 TVTGYTLLLMKYPHVQKVRRELNRGLGAGAPSLGDRTRPLPYTDAVLHQAORLLALVP 368
DB 305 TTLHYGFLMLKYPHYTEKVKQEIQVIGSHRLPTLDDRTKPKYSDAVTHEIQRFSDLLP 364
QY 369 MGIPRLMTTRFRGTYLPCQTEVFPPLGSLHDPNIFKHPBEFNDRFLDADGRPKHE 428
DB 365 IGPHRVTVTKTLFRGYLLPKNTVEYPILSALHDPQYFQPDSPFNPDQFLDANGALKKE 424
QY 429 AFLPF-----SLGKRVCLGEGAKAELFFFTTILQAFSLSPCPDPTLSLKPVTYS 479
DB 425 AFLPSTGQIFQOKSGVKRICIGIESIARSELFFFTSILQNFVSASHVAPKIDILTPKES 484

QY 480 GLFNIPPAFQL 490
DB 485 GIGKIPPTYQI 495

RESULT 11
JT0676

Cytochrome P450 2B - green monkey
C:Species: Cercopithecus aethiops (green monkey, grivet)
C:Date: 18-Nov-1994 #sequence_revision 06-Jan-1995 #text_change 04-Mar-2000
C:Accession: JT0676
R:Ohmori, S.; Sakamoto, Y.; Nakasa, H.; Horie, T.; Saito, K.; Kitada, M.
submitted to JTPID, May 1994
A:Description: Nucleotide and amino acid sequences of monkey P450 2B gene subfamily.
A:Reference number: JT0676
A:Accession: JT0676
A:Molecule type: mRNA
A:Residues: 1-491 <OH>
A:Experimental source: liver
C:Genetics:
A:Gene: CYP2B17
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; heme; iron; metalloprotein; transmembrane protein
F:295-458/Domain: cytochrome P450 homology <P45>
F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 48.8%; Score 1275; DB 2; Length 491;
Best Local Similarity 49.3%; Pred. No. 5.8e-96;
Matches 241; Conservative 100; Mismatches 144; Indels 4; Gaps 3;

QY 8 ALLIALALL--LLLTALSGTRAGHLPDPTPLPILGNLQLRPGALYSGLMRLSKYKG 65
DB 4 SVLLFLALLTGLLLLVQRHFNHAGRLPGCPPLPILGNLQMDRRGLLRFLRFREKYG 63
QY 66 PVFTYLGPRPVVVLVGOEAVREALGGQAEESGRTVAMLEGTGCHGVFNSGGERWQR 125
DB 64 DVTYVILGP-RPVVLLCGVEARREALVDNAEASGRGKIATIDPVFGYGVIFANGNRWK 122
QY 126 QLRKFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPDPSLLLAQAATSNVVCSL 185
DB 123 VLRRFSLTTWRDGMKRSVEERIQEAEQCLVEELRKSQALVDPTFLFHSITANIICSI 182
QY 186 LFGIRSYEDKEFEQAVVRAAGTLLGVSSQGGQTYEMFSLRPLRPGPKQLLHVSTLA 245
DB 183 VFGRHYQDEFLKINLFYHTFSLASSMFGQLFELLSGFLKYFPCAHRQVYKNLQEI 242
QY 246 AFTVROVOQHGNLDASGPARDLVDAFLKMAQEEONPGTEFTKNMMLTVIYLLFAGTM 305
DB 243 AYIGHSVKXRETLDPSAP-QDLIDSYLLQMEKESNPHSEFHSRNLIIINTLSLFFAGTE 301
QY 306 TVSTTVGYTLLLMKYPHVQKVRRELNRGLGAGAPSLGDRTRPLPYTDAVLHQAORLLA 365
DB 302 TTSTTLRYGFLMLKYPHYVAERIYKEIQVIGHRPPALDDRAKMPYTEAVTHEIQRFAD 361
QY 366 LYPMGIPRLMTTRFRGTYLPCQTEVFPPLGSLHDPNIFKHPBEFNDRFLDADGRFR 425
DB 362 LUPMGVPHIVTQTSFRGYIIPKDETEVFELSLDHPHYFEKPDFTFNDPDLANGALK 421
QY 426 KHEAFPLSGLKRVCLGEGAKAELFFFTTILQAFSLSPCPDPTLSLKPVTYSGLFNIP 485
DB 422 KNEAFIPSLGRMCLGEGIARNEUFLFTTILQNFVSAPVEDIDLTPOESGVGKIP 481
QY 486 PAFLQVRP 494
DB 482 PTYQIRFLP 490

RESULT 12

S11305
Cytochrome P450 2B11 - dog
N:Alternate names: cytochrome P450IIB
N:Contains: oxidoreductase (EC 1.1.1.1)
C:Species: Canis lupus familiaris (dog)

C;Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 04-Mar-2000
C;Accession: S11305
R;Graves, P.E.; Elhag, G.A.; Ciaccio, P.J.; Bourque, D.P.; Halpert, J.R.
Arch. Biochem. Biophys. 281, 106-115, 1990
A;Title: cDNA and deduced amino acid sequences of a dog hepatic cytochrome P450IIB respo
F;295-458/Domain: cytochrome P450 homology <P45>
A;Reference number: S11305; MUID:90343348; PMID:2116765
A;Accession: S11305
A;Molecule type: mRNA
A;Residues: 1-494 <GRA>
A;Cross-references: EMBL:M92447; NID:g164039; PIDN:AAA30881.1; PID:g164040
C;Genetics:
A;Gene: CYP2B11
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F;295-458/Domain: cytochrome P450 homology <P45>
F;436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 48.3%; Score 1262; DB 2; Length 494;
Best Local Similarity 49.2%; Pred. No. 6.7e-95;
Matches 237; Conservative 99; Mismatches 144; Indels 2; Gaps 2;
2;
QY 9 LLLALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYGLMLRSKKYGPVF 68
DB 7 LLLALLTGLLLMARGHPKAYGHLPPGPRPLPILGNFLQMDRKGLLKSFLRLQEKYGDVF 66
QY 69 TIYLGPRPVPVVLVGOEAVREALGQAEFSGRGTVAMLEGTDFDGHGVFSGRWRLR 128
DB 67 TVHLGP-RPVVVLCTQTIREALVDHAEAFSGRGTTAAQLVMQDYGIFASGQWKTLR 125
QY 129 KFTMLALRLDGMKREGBELIOAEARCLVETPQGTGRRPDPSPSLIAQAATSNVCSLLFG 188
DB 126 RFLATMRDFGMRKSVVERIOEAEQCLVEELKRYQGVLPDPTFLFQCITANICSIIVFG 185
QY 189 LRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLRPLGPHKQLLHVSTLAFT 248
DB 186 KFGYKDPFELRLMNLFFVFSFALISSFQMFELFSLKYPFGTHRQVYNNLQEIKAPI 245
QY 249 VRQVOHQNLDAAGPARDLVDAFLKMAQEQNPGTFTNKNMLMTVLYLLFAGTMTVS 308
DB 246 ARMVEXHRETLDPASAP-RDFIDAYLRMDKEAAPSSEFHHRLNLDLTSLSLFAGTETS 304
QY 309 TTVGTYLLMLKYPHVKWREELNRELGAQAPSLGDRTRLPYTDVLAHQAQRLALVP 368
DB 305 TTLRYGFLMLKYPHAEIRYKIDQVIGPHRLPSLDDRAKWPYTDVAIHEIQRFGLLP 364
QY 369 MGIPRTLMRTTRPGTYLPQGTVEFPLLSILHDPNIFKHPEFNPDRDLADGRPKHE 428
DB 365 IGVPHVMTKDICFRGYIIPKGTVEFPLLSALNDPHYFEKPDVFNPDHFDANGALKKNE 424
QY 429 AFLPFLSGKVCGLGEGLAKAELEFPTTILQAFSLSPCPDPTLSLKPTVSGLFNPPAF 488
DB 425 AFIPFSGIKRICLGEIARMELFFPTTILQNFVSASPMAPEDIDITPQSIGVGLKPPVY 484
QY 489 QL 490
DB 485 QI 486

RESULT 13
184735
testosterone 16a-hydroxylase type a - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 04-Mar-2000
C;Accession: I84735
R;Lakso, M.; Masaki, R.; Noshiro, M.; Negishi, M.
Eur. J. Biochem. 195, 477-486, 1991
A;Title: Structures and characterization of sex-specific mouse cytochrome P-450 genes as
A;Reference number: I49625; MUID:91146586; PMID:1997326
A;Accession: I84735
A;Status: preliminary; translated from GB/EMBL/DBD
A;Molecule type: DNA
A;Residues: 1-491 <RES>
A;Cross-references: GB:M60273; NID:g192906; PIDN:AAA03648.1; PID:g192908

Query Match 48.3%; Score 1262; DB 2; Length 494;
Best Local Similarity 49.2%; Pred. No. 6.7e-95;
Matches 237; Conservative 99; Mismatches 144; Indels 2; Gaps 2;
2;
QY 9 LLLALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYGLMLRSKKYGPVF 68
DB 7 LLLALLTGLLLMARGHPKAYGHLPPGPRPLPILGNFLQMDRKGLLKSFLRLQEKYGDVF 66
QY 69 TIYLGPRPVPVVLVGOEAVREALGQAEFSGRGTVAMLEGTDFDGHGVFSGRWRLR 128
DB 67 TVHLGP-RPVVVLCTQTIREALVDHAEAFSGRGTTAAQLVMQDYGIFASGQWKTLR 125
QY 129 KFTMLALRLDGMKREGBELIOAEARCLVETPQGTGRRPDPSPSLIAQAATSNVCSLLFG 188
DB 126 RFLATMRDFGMRKSVVERIOEAEQCLVEELKRYQGVLPDPTFLFQCITANICSIIVFG 185
QY 189 LRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLRPLGPHKQLLHVSTLAFT 248
DB 186 KFGYKDPFELRLMNLFFVFSFALISSFQMFELFSLKYPFGTHRQVYNNLQEIKAPI 245
QY 249 VRQVOHQNLDAAGPARDLVDAFLKMAQEQNPGTFTNKNMLMTVLYLLFAGTMTVS 308
DB 246 ARMVEXHRETLDPASAP-RDFIDAYLRMDKEAAPSSEFHHRLNLDLTSLSLFAGTETS 304
QY 309 TTVGTYLLMLKYPHVKWREELNRELGAQAPSLGDRTRLPYTDVLAHQAQRLALVP 368
DB 305 TTLRYGFLMLKYPHAEIRYKIDQVIGPHRLPSLDDRAKWPYTDVAIHEIQRFGLLP 364
QY 369 MGIPRTLMRTTRPGTYLPQGTVEFPLLSILHDPNIFKHPEFNPDRDLADGRPKHE 428
DB 365 IGVPHVMTKDICFRGYIIPKGTVEFPLLSALNDPHYFEKPDVFNPDHFDANGALKKNE 424
QY 429 AFLPFLSGKVCGLGEGLAKAELEFPTTILQAFSLSPCPDPTLSLKPTVSGLFNPPAF 488
DB 425 AFIPFSGIKRICLGEIARMELFFPTTILQNFVSASPMAPEDIDITPQSIGVGLKPPVY 484
QY 489 QL 490
DB 485 QI 486

C;Genetics:
A;Gene: 16aoh-a
A;Introns: 57/3; 112/1; 162/1; 215/3; 274/3; 322/1; 384/3; 432/1
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; heme; iron; metalloprotein; transmembrane protein
F;295-458/Domain: cytochrome P450 homology <P45>
F;436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.7%; Score 1247; DB 2; Length 491;
Best Local Similarity 48.8%; Pred. No. 1.1e-93;
Matches 235; Conservative 99; Mismatches 146; Indels 2; Gaps 2;
2;
QY 9 LLLALLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYGLMLRSKKYGPVF 68
DB 7 LLLALLLSLFLVVRGHAKHGHLPPLPPLGNLLQMDRGGLKLCFQIQERHGDVF 66
QY 69 TIYLGPRPVPVVLVGOEAVREALGQAEFSGRGTVAMLEGTDFDGHGVFSGRWRLR 128
DB 67 TVHLGP-RPVVVLCTQTIREALVDHAEAFSGRGTTAAQLVMQDYGIFASGQWKTLR 125
QY 129 KFTMLALRLDGMKREGBELIOAEARCLVETPQGTGRRPDPSPSLIAQAATSNVCSLLFG 188
DB 126 RFLATMRDFGMRKSVVERIOEAEQCLVEELKRYQGVLPDPTFLFQCITANICSIIVFG 185
QY 189 LRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLRPLGPHKQLLHVSTLAFT 248
DB 186 BRFDYTDQFLHLLMLKIFLSLSFSGQMFELFSGFLKYPFVHRQVIRKQQLLDYI 245
QY 249 VRQVOHQNLDAAGPARDLVDAFLKMAQEQNPGTFTNKNMLMTVLYLLFAGTMTVS 308
DB 246 AHSVEKHATLDPASAP-RDYIDTYLLRMEKESNINTEHHQNLMMVLSLFFAGTETS 304
QY 309 TTVGTYLLMLKYPHVKWREELNRELGAQAPSLGDRTRLPYTDVLAHQAQRLALVP 368
DB 305 ATLHGVLLMLKYPHTEKVKQEIQVIGSHRLPLTDDRDKMPYTDVAIHEIQRFGLLP 364
QY 369 MGIPRTLMRTTRPGTYLPQGTVEFPLLSILHDPNIFKHPEFNPDRDLADGRPKHE 428
DB 365 IGLPHKVIKDTLFRGYLLPKTEVYPLSSALHDPQYFEQPKFNFHFDANGALKKCE 424
QY 429 AFLPFLSGKVCGLGEGLAKAELEFPTTILQAFSLSPCPDPTLSLKPTVSGLFNPPAF 488
DB 425 AFLPFTSGKRICLGEIARNELFFPTTILQNFVSASPMAPKIDITPESGICKIPPAH 484
QY 489 QL 490
DB 485 QI 486

RESULT 14
A31047
testosterone 16alpha-hydroxylase (EC 1.14.14.-) cytochrome P450 2B9 - mouse
N;Alternate names: cytochrome I-P450 16alpha; cytochrome P450pf26
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 04-Mar-2000
C;Accession: A31047
R;Noshiro, M.; Lakso, M.; Kawajiri, K.; Negishi, M.
Biochemistry 27, 6434-6443, 1988
A;Title: Rip locus: regulation of female-specific isozyme (I-P-450-16-alpha) of testosterone
A;Reference number: A31047; MUID:9118235; PMID:3219345
A;Accession: A31047
A;Molecule type: mRNA
A;Residues: 1-491 <NOS>
A;Cross-references: EMBL:M21855; NID:g201966; PIDN:AAA40424.1; PID:g201967
A;Note: the authors translated the codon CAG for residues 105 as Gly, CAG for residue 1
73 as Met, ATG for residue 290 as Asn, GTC for residue 298 as Ala, AAA for residue 323
C;Genetics:
A;Gene: Cyp2b-9
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metal
F;295-458/Domain: cytochrome P450 homology <P45>
F;436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.6%; Score 1244; DB 2; Length 491;
 Best Local Similarity 48.5%; Pred. No. 1.9e-93;
 Matches 234; Conservative 100; Mismatches 146; Indels 2; Gaps 2;

QY 9 LLLALLLLLTALSGTRARGHLPPGTPPLIGNLLQLRPGALYSGLMR 59
 DB 7 LLLAVLLSLFLLVRGHAKIHGHLPPGPHPLFLIGNLLQMDRGLLKCFIQLEKHGDPV 66
 QY 69 TIYLGPRPVPVVLVGOEAVREALGQAEFEFGSGRTVMLEGTDFGHGFFNGERWRQIR 128
 DB 67 TVHLGP-RPVPVVLGCTQIRREALVDHAFAFGSGRTIAAAQVLVMQDYGIFFASGQRWKTLR 125
 QY 129 KFTMLARDLGMKREGEELIOAEARCLVETFGTGEGRFPDFSLLLAQATSNVVCSSLF 188
 DB 126 RFLSLATMKEFGMGKRSVEERIKEEAQCLVEELKQYGVPLDPTFLFCQITANIICSVIF 185
 QY 189 LRFSEYDEKEFOAVRAAGTLLGVSSQGGQYEMFSLRPLDPGPHKQLLHHVSTLAFT 248
 DB 186 ERFDYTDQFHLNLMYKIFSLSSFSFGQFELFSGFLKYFPGVHRQIVKKQQLLDYI 245
 QY 249 VRQVQHQGNLDASGPARDLVDAFLLKMAQEEQNGTFTNKNMLMTVIYLLFAGTMTVS 308
 DB 246 AHSVEKHKATLDSPAP-RDVIDTYLLRMEKEKSHNTEFFHQNLMMSVLSFFVGTETTS 304
 QY 309 TTGVYTLILLMKYPHVQKWRBELNRELGAQAPSLGDRTRLPYTDVHLHEAQQLLALVP 368
 DB 305 ATLLHYGVLLMLKYPHVTEKQKEIDQVIGSHRLPTLDDRTKMPYTDVAIHEIQRFSDLP 364
 QY 369 MGIPRTLMRTTRRGYTLPGQTEVPFLLGSILHDPNIFKPEEPNDRFIDADGRFRKHE 428
 DB 365 IGLPHKVIKDTLFRGYLLLPKNTEVYVLSALHDPQYFEQDFKFNPEQFLDANGALKKE 424
 QY 429 AFLPFLSGKRVCLGEGLAKEALFFFTTILQAFSLSPCPDPTLSLKPVTSGLENIPAP 488
 DB 425 AFLPFSGKRICLGESIAARNELFIFFTTILQNFVSASVPAPKIDIDLTPKSGIGKIPPAH 484
 QY 489 QL 490
 DB 485 QI 486

RESULT 15
 A3551
 cytochrome P450 2G1 - rat
 N:Contains: oxidoreductase (EC 1.-.-.-)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 04-Mar-2000
 C:Accession: A35551, A33875
 R:Nef, P.; Larabee, T.M.; Kagimoto, K.; Meyer, U.A.
 J. Biol. Chem. 265, 2903-2907, 1990
 A>Title: Olfactory-specific cytochrome P-450 (P-450olf1; IIG1). Gene structure and devel
 A:Reference number: A35551, MUID:90153922; PMID:2406242
 A:Accession: A35551
 A:Molecule type: DNA
 A:Residues: 1-494 <NEF>
 A:Cross-references: GB:M31931; GB:J04715; NID:g203888; PIDN:AAA41069.1; PID:g203890
 R:Nef, P.; Heldman, J.; Lazar, D.; Margalit, T.; Jaye, M.; Hanukoglu, I.; Lancet, D.
 J. Biol. Chem. 264, 6780-6785, 1989
 A>Title: Olfactory-specific cytochrome P-450. cDNA cloning of a novel neuroepithelial en
 A:Reference number: A33875; MUID:89214086; PMID:2708343
 A:Accession: A33875
 A:Molecule type: mRNA
 A:Residues: 2-494 <NE2>
 A:Cross-references: GB:M31931; GB:J04715; NID:g203888; PIDN:AAA41069.1; PID:g203890
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;
 F:298-461/Domain: cytochrome P450 homology <P45>
 F:439/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.6%; Score 1243.5; DB 2; Length 494;
 Best Local Similarity 47.7%; Pred. No. 2.1e-93;
 Matches 234; Conservative 101; Mismatches 153; Indels 3; Gaps 3;

QY 1 MEATGTWALLALAL-LLLLLTIALSGTRARGHLPPGTPPLIGNLLQLRPGALYSGLMR 59
 DB 1 MALGGAFFSIFMTCLSCILLIAMIWKRTSRGGKLPDGPPIPIPLIGNLLQVRIDATQSFLLK 60
 QY 60 LSKKYGPVFTIYLGPRPVPVVLVGOEAVREALGQAEFEFGSGRTVMLEGTDFGHGVPFS 119
 DB 61 LQKYGSVFTYVFGP-RPVPVVLGCTQIRREALVDHAFAFGSGRTIAAAQVLVMQDYGIFFASG 119
 QY 120 NGERWRQLRKFETMLARDLGMKREGEELIOAEARCLVETFGTGEGRFPDFSLLLAQATS 179
 DB 120 NGERWKILRRSLVLRNFGMGKRSIERIOEAGYLLLEELHKVKGADIDPTFYLSRVTS 179
 QY 180 NVVCSLLFGLRFSYEDKEFOAVRAAGTLLGVSSQGGQYEMFSLRPLDPGPHKQLLH 239
 DB 180 NVICSVVFGKRFYEDQDRFSLMKMINESFVEMSMWMAQLYDMVYGVIOYFPPGRHRLYN 239
 QY 240 HVSTLAAFTVRQVQHQGNLDASGPARDLVDAFLLKMAQEEQNGTFTNKNMLMTVIY 299
 DB 240 LIEELKDFIASRVKINEASDPSPN-RDFIDCFILKMYQDKSDPHSEFNKMLVLTLLN 298
 QY 300 LFACTMTVSTTVGYTLTLLMKYPHVQKWRBELNRELGAQAPSLGDRTRLPYTDVHLHE 359
 DB 299 FFAGTETVSTLRYGFLLLMKYPEVEAKIHEEINQVIGTHRTTPVDDRAKMPYTDVHLHE 358
 QY 360 AORLLALVPMGIPRTLMRTTRRGYTLPGQTEVPFLLGSILHDPNIFKPEEPNDRF 419
 DB 359 IQRLTDIVPLGVPHNVIRDTFRGYFLPKGTVVYVPLIGSVLKDQPKYFRYPEAFYPQHFLD 418
 QY 420 ADGRFRKHEAFLPFLSGKRVCLGEGLAKEALFFFTTILQAFSLSPCPDPTLSLKPVT 479
 DB 419 EQGRFKNDFAVAFSSGKRICVGEALAMELEFLYFTSILQRFSLRSLVPPADIDIAHKIS 478
 QY 480 GLFNIPAPOL 490
 DB 479 GFGNIPETVEL 489

Search completed: September 15, 2004, 09:21:56
 Job time : 45 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:21:16 ; Search time 128 Seconds

(without alignments)
1262.715 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	2615	100.0	504	9	US-09-748-127-2
2	2615	100.0	504	12	US-10-147-493-64
3	2615	100.0	504	12	US-10-258-080-6
4	2615	100.0	504	12	US-10-145-127-64
5	2615	100.0	504	12	US-10-160-503-64
6	2615	100.0	504	12	US-10-143-118-64
7	2615	100.0	504	12	US-10-144-993-64
8	2615	100.0	504	12	US-10-158-787-64
9	2615	100.0	504	12	US-10-669-693-2
10	2615	100.0	504	12	US-10-140-024-64
11	2615	100.0	504	12	US-10-140-808-64
12	2615	100.0	504	12	US-10-152-405-64
13	2615	100.0	504	12	US-10-127-852A-64
14	2615	100.0	504	12	US-10-127-900A-64
15	2615	100.0	504	12	US-10-128-685A-64

16	2615	100.0	504	12	US-10-131-820A-64	Sequence 64, Appl
17	2615	100.0	504	12	US-10-142-886-64	Sequence 64, Appl
18	2615	100.0	504	12	US-10-146-728-64	Sequence 64, Appl
19	2615	100.0	504	12	US-10-146-786-64	Sequence 64, Appl
20	2615	100.0	504	12	US-10-147-499-64	Sequence 64, Appl
21	2615	100.0	504	12	US-10-157-798-64	Sequence 64, Appl
22	2615	100.0	504	14	US-10-028-072-64	Sequence 5, Appl
23	2615	100.0	504	14	US-10-121-049-64	Sequence 64, Appl
24	2615	100.0	504	14	US-10-123-904-64	Sequence 64, Appl
25	2615	100.0	504	14	US-10-140-470-64	Sequence 64, Appl
26	2615	100.0	504	14	US-10-067-668-5	Sequence 64, Appl
27	2615	100.0	504	14	US-10-175-746-64	Sequence 64, Appl
28	2615	100.0	504	14	US-10-176-918-64	Sequence 64, Appl
29	2615	100.0	504	14	US-10-176-921-64	Sequence 64, Appl
30	2615	100.0	504	14	US-10-137-865-64	Sequence 64, Appl
31	2615	100.0	504	14	US-10-140-474-64	Sequence 64, Appl
32	2615	100.0	504	14	US-10-142-431-64	Sequence 64, Appl
33	2615	100.0	504	14	US-10-143-114-64	Sequence 64, Appl
34	2615	100.0	504	14	US-10-140-002-64	Sequence 64, Appl
35	2615	100.0	504	14	US-10-142-419-64	Sequence 64, Appl
36	2615	100.0	504	14	US-10-123-262-64	Sequence 64, Appl
37	2615	100.0	504	14	US-10-142-423-64	Sequence 64, Appl
38	2615	100.0	504	14	US-10-121-050-64	Sequence 64, Appl
39	2615	100.0	504	14	US-10-141-755-64	Sequence 64, Appl
40	2615	100.0	504	14	US-10-143-032-64	Sequence 64, Appl
41	2615	100.0	504	14	US-10-123-108-64	Sequence 64, Appl
42	2615	100.0	504	14	US-10-123-236-64	Sequence 64, Appl
43	2615	100.0	504	14	US-10-123-261-64	Sequence 64, Appl
44	2615	100.0	504	14	US-10-140-921-64	Sequence 64, Appl
45	2615	100.0	504	14	US-10-140-928-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-09-748-127-2
; Sequence 2, Application US/09748127
; Patent No. US2002007674A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000685
; CURRENT APPLICATION NUMBER: US/09/748,127
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Human
US-09-748-127-2

Query Match	100.0%;	Score 2615;	DB 9;	Length 504;
Best Local Similarity	100.0%;	Pred. No. 8.1e-253;		
Matches 504;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MEATGTWALLALALLLLTALSGTRARGHLP	PGTPTPLLGNNLLQLRFGALYGLMRL	60
Db	1	MEATGTWALLALALLLLTALSGTRARGHLP	PGTPTPLLGNNLLQLRFGALYGLMRL	60
QY	61	SKKYGVFTYILGPRPVVVLVGQAVREALGQAE	EEFSGRGTVMLEGTFDGCVFFSN	120
Db	61	SKKYGVFTYILGPRPVVVLVGQAVREALGQAE	EEFSGRGTVMLEGTFDGCVFFSN	120
QY	121	GERWQLRKFTMLALDLGMKREGEELIQARCLV	ETFGTEGPRPDPSSLLLAQATSN	180
Db	121	GERWQLRKFTMLALDLGMKREGEELIQARCLV	ETFGTEGPRPDPSSLLLAQATSN	180
QY	181	VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGV	SSQGGQTYEMFSWFLRPLPGFKQLLHH	240

Db	121	GERWRQLRKFMTLALRDLGMGKREGEELIQAEARCLVETFGTEGRFPDPSLLLAQATSN	180
Qy	181	VVCILLGLRFSYEDKQFQAVVRAAGGTLGVSQGGQTYEMFSLRPLPGPHKQLLHH	240
Db	181	VVCILLGLRFSYEDKQFQAVVRAAGGTLGVSQGGQTYEMFSLRPLPGPHKQLLHH	240
Qy	241	VSTLAAFTVRQVQOQHGNLDASGPARDLDVDAFLKMAQEQNPGTEFTNKMLMTVIYLL	300
Db	241	VSTLAAFTVRQVQOQHGNLDASGPARDLDVDAFLKMAQEQNPGTEFTNKMLMTVIYLL	300
Qy	301	FAGTMVSTTVGYTLLLLMKYPHVQKVVREELNRELGLGAGAPSLGDRTRLPTDVLHEA	360
Db	301	FAGTMVSTTVGYTLLLLMKYPHVQKVVREELNRELGLGAGAPSLGDRTRLPTDVLHEA	360
Qy	361	ORLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPPLGSLILHDPNTFKHPEENPDRFLDA	420
Db	361	ORLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPPLGSLILHDPNTFKHPEENPDRFLDA	420
Qy	421	DGRFRKHEAFPLFSLGKRVCLGEGGLAKAELFLFFTTILQAFSLESFPCPPDTSLSKPTVSG	480
Db	421	DGRFRKHEAFPLFSLGKRVCLGEGGLAKAELFLFFTTILQAFSLESFPCPPDTSLSKPTVSG	480
Qy	481	LFNIPPAFQQLQVRPTDLHSTTQTR	504
Db	481	LFNIPPAFQQLQVRPTDLHSTTQTR	504

RESULT 3

US-10-147-493-64

Sequence 64, Application US/10147493

Publication No. US20040029217A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: F330R1C345

CURRENT APPLICATION NUMBER: US/10/147,493

CURRENT FILING DATE: 2002-05-17

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 64

LENGTH: 504

TYPE: PRT

ORGANISM: Homo Sapien

US-10-147-493-64

Query Match

Best Local Similarity

Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

100.0%; Score 2615; DB 12; Length 504;

100.0%; Pred. No. 8.1e-253;

Query Match

Best Local Similarity

Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

100.0%; Score 2615; DB 12; Length 504;

100.0%; Pred. No. 8.1e-253;

Db	181	VVCILLGLRFSYEDKQFQAVVRAAGGTLGVSQGGQTYEMFSLRPLPGPHKQLLHH	240
Qy	241	VSTLAAFTVRQVQOQHGNLDASGPARDLDVDAFLKMAQEQNPGTEFTNKMLMTVIYLL	300
Db	241	VSTLAAFTVRQVQOQHGNLDASGPARDLDVDAFLKMAQEQNPGTEFTNKMLMTVIYLL	300
Qy	301	FAGTMVSTTVGYTLLLLMKYPHVQKVVREELNRELGLGAGAPSLGDRTRLPTDVLHEA	360
Db	301	FAGTMVSTTVGYTLLLLMKYPHVQKVVREELNRELGLGAGAPSLGDRTRLPTDVLHEA	360
Qy	361	ORLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPPLGSLILHDPNTFKHPEENPDRFLDA	420
Db	361	ORLLALVPMGIPRTLMRTTRFRGYTLPGQTEVPPLGSLILHDPNTFKHPEENPDRFLDA	420
Qy	421	DGRFRKHEAFPLFSLGKRVCLGEGGLAKAELFLFFTTILQAFSLESFPCPPDTSLSKPTVSG	480
Db	421	DGRFRKHEAFPLFSLGKRVCLGEGGLAKAELFLFFTTILQAFSLESFPCPPDTSLSKPTVSG	480
Qy	481	LFNIPPAFQQLQVRPTDLHSTTQTR	504
Db	481	LFNIPPAFQQLQVRPTDLHSTTQTR	504

RESULT 2

US-10-147-493-64

Sequence 64, Application US/10147493

Publication No. US20040029217A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: F330R1C345

CURRENT APPLICATION NUMBER: US/10/147,493

CURRENT FILING DATE: 2002-05-17

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 64

LENGTH: 504

TYPE: PRT

ORGANISM: Homo Sapien

US-10-147-493-64

Query Match

Best Local Similarity

Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

100.0%; Score 2615; DB 12; Length 504;

100.0%; Pred. No. 8.1e-253;

Query Match

Best Local Similarity

Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

100.0%; Score 2615; DB 12; Length 504;

100.0%; Pred. No. 8.1e-253;

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040029125A1 1272843CD1
US-10-258-080-6

Query Match      100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTALSCTRARGHLPPGTPPLLLGNLLQRLRGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSCTRARGHLPPGTPPLLLGNLLQRLRGALYSGLMRL 60
QY 61 SKYGVFTIYLGPMRPVVVLVQAEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120
Db 61 SKYGVFTIYLGPMRPVVVLVQAEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120
QY 121 GERWRQLRKFTMLALRDGLMGKEGEBELIOAEARCLVETFGTEGPRPDPSSLQAQTSN 180
Db 121 GERWRQLRKFTMLALRDGLMGKEGEBELIOAEARCLVETFGTEGPRPDPSSLQAQTSN 180
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAFTVRQVQOHOGLNDASGPARDLVDAFLLLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
Db 241 VSTLAFTVRQVQOHOGLNDASGPARDLVDAFLLLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
QY 301 FACTMTVSTTVGTYTLMLMKYPHVQKWVRELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
Db 301 FACTMTVSTTVGTYTLMLMKYPHVQKWVRELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
QY 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSIILHDNIFKHPEEFPDRFLDA 420
Db 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSIILHDNIFKHPEEFPDRFLDA 420
QY 421 DGRFRKHEAPFLPSLGRVCLGEGAKAEFLPFTTILQAFSLESPPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAPFLPSLGRVCLGEGAKAEFLPFTTILQAFSLESPPDPTLSLKPTVSG 480
QY 481 LFNIPPAFQLOVRPTDLHSTTQTR 504
Db 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

RESULT 4
US-10-145-127-64
; Sequence 64, Application US/10145127
; Publication No. US20040033558A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C252
; CURRENT APPLICATION NUMBER: US/10/145,127
```

```
; CURRENT FILING DATE: 2002-05-13
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-145-127-64

Query Match      100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTALSCTRARGHLPPGTPPLLLGNLLQRLRGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSCTRARGHLPPGTPPLLLGNLLQRLRGALYSGLMRL 60
QY 61 SKYGVFTIYLGPMRPVVVLVQAEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120
Db 61 SKYGVFTIYLGPMRPVVVLVQAEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120
QY 121 GERWRQLRKFTMLALRDGLMGKEGEBELIOAEARCLVETFGTEGPRPDPSSLQAQTSN 180
Db 121 GERWRQLRKFTMLALRDGLMGKEGEBELIOAEARCLVETFGTEGPRPDPSSLQAQTSN 180
QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAFTVRQVQOHOGLNDASGPARDLVDAFLLLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
Db 241 VSTLAFTVRQVQOHOGLNDASGPARDLVDAFLLLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
QY 301 FACTMTVSTTVGTYTLMLMKYPHVQKWVRELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
Db 301 FACTMTVSTTVGTYTLMLMKYPHVQKWVRELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
QY 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSIILHDNIFKHPEEFPDRFLDA 420
Db 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSIILHDNIFKHPEEFPDRFLDA 420
QY 421 DGRFRKHEAPFLPSLGRVCLGEGAKAEFLPFTTILQAFSLESPPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAPFLPSLGRVCLGEGAKAEFLPFTTILQAFSLESPPDPTLSLKPTVSG 480
QY 481 LFNIPPAFQLOVRPTDLHSTTQTR 504
Db 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

RESULT 5
US-10-160-503-64
; Sequence 64, Application US/10160503
; Publication No. US20040033559A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C446
; CURRENT APPLICATION NUMBER: US/10/160,503
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-160-503-64

Query Match 100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTLALSCTRARGHLPPOGTPPLPLGNLLQLRPGALYGLMRL 60
DB 1 MEATGTWALLALALLLLTLALSCTRARGHLPPOGTPPLPLGNLLQLRPGALYGLMRL 60
QY 61 SKKYGPVFTIYLGPNRPVVVLVGQAVREALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120
DB 61 SKKYGPVFTIYLGPNRPVVVLVGQAVREALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120
QY 121 GERWQLRKFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPFDPSSLILAQATSN 180
DB 121 GERWQLRKFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPFDPSSLILAQATSN 180
QY 181 VVCSLLFGLRSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLPGPHKQLLHH 240
QY 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPGETFNKNMLMTVIYLL 300
DB 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPGETFNKNMLMTVIYLL 300
QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKVRBELNRELGLGAGQPSLDGRTLRPYTDAVLHEA 360
DB 301 FAGTMTVSTTVGYTLLLMKYPHVQKVRBELNRELGLGAGQPSLDGRTLRPYTDAVLHEA 360
QY 361 QRLALVPNGIPRTLMRTTRFRGYTLPGQTEVFPPLGSLILHDPNIFKPIPEENPDRFLDA 420
DB 361 QRLALVPNGIPRTLMRTTRFRGYTLPGQTEVFPPLGSLILHDPNIFKPIPEENPDRFLDA 420
QY 421 DGRFRKHEAFLPFSLGKRVCLGEGAKAELFLFTTILQAFLSPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFSLGKRVCLGEGAKAELFLFTTILQAFLSPCPDPTLSLKPTVSG 480
QY 481 LFNIPPAFQLOVRPTDLHSTTQTR 504
DB 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

RESULT 6

US-10-143-118-64
; Sequence 64, Application US/10143118
; Publication No. US20040038335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C228
; CURRENT APPLICATION NUMBER: US/10/143,118
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-143-118-64

Query Match 100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTLALSCTRARGHLPPOGTPPLPLGNLLQLRPGALYGLMRL 60
DB 1 MEATGTWALLALALLLLTLALSCTRARGHLPPOGTPPLPLGNLLQLRPGALYGLMRL 60
QY 61 SKKYGPVFTIYLGPNRPVVVLVGQAVREALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120
DB 61 SKKYGPVFTIYLGPNRPVVVLVGQAVREALGQAEFSGRGTVAMLEGTDFDGHGVFFSN 120
QY 121 GERWQLRKFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPFDPSSLILAQATSN 180
DB 121 GERWQLRKFTMLALRDLGMKREGEELIQAEARCLVETFOGTEGRPFDPSSLILAQATSN 180
QY 181 VVCSLLFGLRSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSLRPLPGPHKQLLHH 240
QY 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPGETFNKNMLMTVIYLL 300
DB 241 VSTLAAFTVRQVQOQHGNDASGPARDLVDAFLKMAQEEQNPGETFNKNMLMTVIYLL 300
QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKVRBELNRELGLGAGQPSLDGRTLRPYTDAVLHEA 360
DB 301 FAGTMTVSTTVGYTLLLMKYPHVQKVRBELNRELGLGAGQPSLDGRTLRPYTDAVLHEA 360
QY 361 QRLALVPNGIPRTLMRTTRFRGYTLPGQTEVFPPLGSLILHDPNIFKPIPEENPDRFLDA 420
DB 361 QRLALVPNGIPRTLMRTTRFRGYTLPGQTEVFPPLGSLILHDPNIFKPIPEENPDRFLDA 420
QY 421 DGRFRKHEAFLPFSLGKRVCLGEGAKAELFLFTTILQAFLSPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFSLGKRVCLGEGAKAELFLFTTILQAFLSPCPDPTLSLKPTVSG 480
QY 481 LFNIPPAFQLOVRPTDLHSTTQTR 504
DB 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

RESULT 7

US-10-144-993-64
; Sequence 64, Application US/10144993
; Publication No. US20040038336A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.

APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C261
CURRENT APPLICATION NUMBER: US/10/144,993
Prior Filing Date: 2002-05-13
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 64
LENGTH: 504
TYPE: PRT
ORGANISM: Homo Sapien
US-10-144-993-64

Query Match 100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEATGTWALLALALLLLTALSGTRAGHLPQGTPLPLGNLLQRLPGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSGTRAGHLPQGTPLPLGNLLQRLPGALYSGLMRL 60
QY 61 SKKYGVPVFTIYLGWPWPVVLVQAEAVREALGQAEFEFSGRTVAMLEGTDFGHGVFFSN 120
Db 61 SKKYGVPVFTIYLGWPWPVVLVQAEAVREALGQAEFEFSGRTVAMLEGTDFGHGVFFSN 120
QY 121 GERWRQLRKFTMLALRDLMGKREGELIQAEARCLVETFGTEGRPDPSLLLAQATSN 180
Db 121 GERWRQLRKFTMLALRDLMGKREGELIQAEARCLVETFGTEGRPDPSLLLAQATSN 180
QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAAFTVRQVQOQHNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKMLMTVIYLL 300
Db 241 VSTLAAFTVRQVQOQHNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKMLMTVIYLL 300
QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVRELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
Db 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVRELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
QY 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLIHDNIFKHPEEFNPDRLDA 420
Db 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLIHDNIFKHPEEFNPDRLDA 420
QY 421 DGRFRKHEAFPLPSLGRVCLGEGAKAELFLPFTTILQAFSLSPCPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAFPLPSLGRVCLGEGAKAELFLPFTTILQAFSLSPCPDPTLSLKPTVSG 480
QY 481 LFNIPAFQLQVRPTDLHSTTQTR 504
Db 481 LFNIPAFQLQVRPTDLHSTTQTR 504

RESULT 8

US-10-158-787-64
Sequence 64, Application US/10158787
Publication No. US20040039164A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C449
CURRENT APPLICATION NUMBER: US/10/158,787
CURRENT Filing Date: 2003-04-03
Prior Application Number: 60/049911
Prior Filing Date: 1997-06-18
Prior Application Number: 60/056974
Prior Filing Date: 1997-08-26
Prior Application Number: 60/059113
Prior Filing Date: 1997-09-17
Prior Application Number: 60/059115
Prior Filing Date: 1997-09-17
Prior Application Number: 60/059117
Prior Filing Date: 1997-09-17
Prior Application Number: 60/059122
Prior Filing Date: 1997-09-17
Prior Application Number: 60/059184
Prior Filing Date: 1997-09-17
Prior Application Number: 60/059263
Prior Filing Date: 1997-09-18
Prior Application Number: 60/059352
Prior Filing Date: 1997-09-19
Prior Application Number: 60/059588
Prior Filing Date: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 64
LENGTH: 504
TYPE: PRT
ORGANISM: Homo Sapien
US-10-158-787-64

Query Match 100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEATGTWALLALALLLLTALSGTRAGHLPQGTPLPLGNLLQRLPGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSGTRAGHLPQGTPLPLGNLLQRLPGALYSGLMRL 60
QY 61 SKKYGVPVFTIYLGWPWPVVLVQAEAVREALGQAEFEFSGRTVAMLEGTDFGHGVFFSN 120
Db 61 SKKYGVPVFTIYLGWPWPVVLVQAEAVREALGQAEFEFSGRTVAMLEGTDFGHGVFFSN 120
QY 121 GERWRQLRKFTMLALRDLMGKREGELIQAEARCLVETFGTEGRPDPSLLLAQATSN 180
Db 121 GERWRQLRKFTMLALRDLMGKREGELIQAEARCLVETFGTEGRPDPSLLLAQATSN 180
QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAAFTVRQVQOQHNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKMLMTVIYLL 300
Db 241 VSTLAAFTVRQVQOQHNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKMLMTVIYLL 300
QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVRELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
Db 301 FAGTMTVSTTVGYTLLLMKYPHVQKWVRELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
QY 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLIHDNIFKHPEEFNPDRLDA 420
Db 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVPFLGSLIHDNIFKHPEEFNPDRLDA 420
QY 421 DGRFRKHEAFPLPSLGRVCLGEGAKAELFLPFTTILQAFSLSPCPDPTLSLKPTVSG 480

Db 421 DGRFRKHEAFLPFLSLGKRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480
Qy 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504
Db 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504

RESULT 9
US-10-669-693-2
; Sequence 2, Application US/10669693
; Publication No. US20040043413A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000685CON
; CURRENT APPLICATION NUMBER: US/10/669,693
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/748,127
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-669-693-2

Query Match 100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLLLGNLLQRLPGALYGLMRL 60
Db 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLLLGNLLQRLPGALYGLMRL 60

Qy 61 SKKYGVPFTIYLGPMRPVVVLVGQEA VREALGQQAEEFSGRGTVAMLEGTDFDGHGVFFSN 120
Db 61 SKKYGVPFTIYLGPMRPVVVLVGQEA VREALGQQAEEFSGRGTVAMLEGTDFDGHGVFFSN 120

Qy 121 GERWQRLKFTMLALDLGKKEGEBELIQAEARCLVETFOGTEGRFPDPSLLLAQATSN 180
Db 121 GERWQRLKFTMLALDLGKKEGEBELIQAEARCLVETFOGTEGRFPDPSLLLAQATSN 180

Qy 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

Qy 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLLKMAQEQNPGTEFTNKMLMTVIYLL 300
Db 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLLKMAQEQNPGTEFTNKMLMTVIYLL 300

Qy 301 FAGTMTVSTTVGYTLLMLKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDVAVLHEA 360
Db 301 FAGTMTVSTTVGYTLLMLKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDVAVLHEA 360

Qy 361 QRLALVPMGIPRTLMTTRFRGYTLPQGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420
Db 361 QRLALVPMGIPRTLMTTRFRGYTLPQGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420

Qy 421 DGRFRKHEAFLPFLSLGKRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAFLPFLSLGKRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480

Qy 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504
Db 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504

RESULT 10

US-10-140-024-64
; Sequence 64, Application US/10140024
; Publication No. US20040058424A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Bersini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC69
; CURRENT APPLICATION NUMBER: US/10/140,024
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-024-64

Query Match 100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLLLGNLLQRLPGALYGLMRL 60
Db 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLLLGNLLQRLPGALYGLMRL 60

Qy 61 SKKYGVPFTIYLGPMRPVVVLVGQEA VREALGQQAEEFSGRGTVAMLEGTDFDGHGVFFSN 120
Db 61 SKKYGVPFTIYLGPMRPVVVLVGQEA VREALGQQAEEFSGRGTVAMLEGTDFDGHGVFFSN 120

Qy 121 GERWQRLKFTMLALDLGKKEGEBELIQAEARCLVETFOGTEGRFPDPSLLLAQATSN 180
Db 121 GERWQRLKFTMLALDLGKKEGEBELIQAEARCLVETFOGTEGRFPDPSLLLAQATSN 180

Qy 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

Qy 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLLKMAQEQNPGTEFTNKMLMTVIYLL 300
Db 241 VSTLAFTVRQVOHQGNLDASGPARDLVDAFLLKMAQEQNPGTEFTNKMLMTVIYLL 300

Qy 301 FAGTMTVSTTVGYTLLMLKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDVAVLHEA 360
Db 301 FAGTMTVSTTVGYTLLMLKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDVAVLHEA 360

Qy 361 QRLALVPMGIPRTLMTTRFRGYTLPQGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420
Db 361 QRLALVPMGIPRTLMTTRFRGYTLPQGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420

Qy 421 DGRFRKHEAFLPFLSLGKRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAFLPFLSLGKRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480

Qy 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504
Db 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504

RESULT 11
US-10-140-808-64
; Sequence 64, Application US/10140808
; Publication No. US20030017563A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Steward, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C182
; CURRENT APPLICATION NUMBER: US/10/140,808
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-808-64

Query Match 100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEATGTWALLALALLLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60
DB 1 MEATGTWALLALALLLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60
QY 61 SKKYGPFVTIYLGPRPVPVVLVGOEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSN 120
DB 61 SKKYGPFVTIYLGPRPVPVVLVGOEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSN 120
QY 121 GERWRQLRPTMLALRDLMGKREGBELIQAEARCLVETPQGTGPRPDPSSLILAQATSN 180
DB 121 GERWRQLRPTMLALRDLMGKREGBELIQAEARCLVETPQGTGPRPDPSSLILAQATSN 180
QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAAFTVROVQOHOGLNDASGPARDLVDAFLKMAQEONPGTFTNKNMLMTVIYLL 300
DB 241 VSTLAAFTVROVQOHOGLNDASGPARDLVDAFLKMAQEONPGTFTNKNMLMTVIYLL 300
QY 301 FAGTMTVSTTVGTYTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
DB 301 FAGTMTVSTTVGTYTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
QY 361 ORLLALVPMGIPRTLMTTRFRGYTLPGQTEVPPLILGSLIHDNPIFKHPEENPDRFLDA 420
DB 361 ORLLALVPMGIPRTLMTTRFRGYTLPGQTEVPPLILGSLIHDNPIFKHPEENPDRFLDA 420
QY 421 DGRFRKHEAFLPFLSKGRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAFLPFLSKGRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480

QY 481 LENIPPAFOQLVRPTDLHSTTQTR 504
DB 481 LENIPPAFOQLVRPTDLHSTTQTR 504

RESULT 12

US-10-152-405-64
; Sequence 64, Application US/10152405
; Publication No. US20030211571A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Steward, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C383
; CURRENT APPLICATION NUMBER: US/10/152,405
; CURRENT FILING DATE: 2002-05-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-152-405-64

Query Match 100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEATGTWALLALALLLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60
DB 1 MEATGTWALLALALLLLLTALSGTRARGHLPPTPLPLGNLLQLRPGALYSGLMRL 60
QY 61 SKKYGPFVTIYLGPRPVPVVLVGOEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSN 120
DB 61 SKKYGPFVTIYLGPRPVPVVLVGOEAVREALGQAEFFSGRTVAMLEGTFDGHGVFFSN 120
QY 121 GERWRQLRPTMLALRDLMGKREGBELIQAEARCLVETPQGTGPRPDPSSLILAQATSN 180
DB 121 GERWRQLRPTMLALRDLMGKREGBELIQAEARCLVETPQGTGPRPDPSSLILAQATSN 180
QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAAFTVROVQOHOGLNDASGPARDLVDAFLKMAQEONPGTFTNKNMLMTVIYLL 300
DB 241 VSTLAAFTVROVQOHOGLNDASGPARDLVDAFLKMAQEONPGTFTNKNMLMTVIYLL 300
QY 301 FAGTMTVSTTVGTYTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
DB 301 FAGTMTVSTTVGTYTLLLMKYPHVQKWVREELNRELGAQAPSLGDRTRLPYTDVAVLHEA 360
QY 361 ORLLALVPMGIPRTLMTTRFRGYTLPGQTEVPPLILGSLIHDNPIFKHPEENPDRFLDA 420
DB 361 ORLLALVPMGIPRTLMTTRFRGYTLPGQTEVPPLILGSLIHDNPIFKHPEENPDRFLDA 420
QY 421 DGRFRKHEAFLPFLSKGRVCLGEGAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480

Db 61 SKYGPVFTIYLGPRPFWVLVGOEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120
QY 121 GERWRQKFTMLALRDLGMGKREGELIQAEARCLVETFOCTEGRPDPSLLLAQATSN 180
Db 121 GERWRQKFTMLALRDLGMGKREGELIQAEARCLVETFOCTEGRPDPSLLLAQATSN 180
QY 181 VVCSILLFGLRFSYEDKEFQAVVRAAGTTLGVSSGGQGYEMFSLPLPGPHKQLLHH 240
Db 181 VVCSILLFGLRFSYEDKEFQAVVRAAGTTLGVSSGGQGYEMFSLPLPGPHKQLLHH 240
QY 241 VSTLAFTVRQVQOQHGNLDASGPARDLVDAFLMKMAQEQNPGTEFTNKMLMTVIYLL 300
Db 241 VSTLAFTVRQVQOQHGNLDASGPARDLVDAFLMKMAQEQNPGTEFTNKMLMTVIYLL 300
QY 301 FAGTMVSTTVGYTLLMLMKYPHVQKVRRELNRELGLAGQAPSLGDRTRLPYTDVLAHEA 360
Db 301 FAGTMVSTTVGYTLLMLMKYPHVQKVRRELNRELGLAGQAPSLGDRTRLPYTDVLAHEA 360
QY 361 QRLIALVPMGIPRTLMRTTRFRGYTLPGQTEVFFLLGSLHDPNIFKHPEENPDRFLDA 420
Db 361 QRLIALVPMGIPRTLMRTTRFRGYTLPGQTEVFFLLGSLHDPNIFKHPEENPDRFLDA 420
QY 421 DGRFRKHEAFLPFSGLKRVCLGEGAKAELFFFTTILQAFSLBSPCPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAFLPFSGLKRVCLGEGAKAELFFFTTILQAFSLBSPCPDPTLSLKPTVSG 480
QY 481 LFNIPPAFQLOVRPTDLHSTTQTR 504
Db 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

RESULT 14
US-10-127-900A-64
; Sequence 64, Application US/10127900A
; Publication No. US20030203429A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C81
; CURRENT APPLICATION NUMBER: US/10/127,900A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263

Db 421 DGRFRKHEAFLPFSGLKRVCLGEGAKAELFFFTTILQAFSLBSPCPDPTLSLKPTVSG 480
QY 481 LFNIPPAFQLOVRPTDLHSTTQTR 504
Db 481 LFNIPPAFQLOVRPTDLHSTTQTR 504
RESULT 13
US-10-127-852A-64
; Sequence 64, Application US/10127852A
; Publication No. US20030203428A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C88
; CURRENT APPLICATION NUMBER: US/10/127,852A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 64
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-127-852A-64

Query Match 100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEATGTWALLIALALLLLTIALSGTRARGHLPGPPTPLPLGNLQLRPGALYSGLMRL 60
Db 1 MEATGTWALLIALALLLLTIALSGTRARGHLPGPPTPLPLGNLQLRPGALYSGLMRL 60
QY 61 SKYGPVFTIYLGPRPFWVLVGOEAVREALGQAEFFSGRGTVAMLEGTDFGHGVFFSN 120


```
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059352
/ PRIOR FILING DATE: 1997-09-19
/ PRIOR APPLICATION NUMBER: 60/059588
/ PRIOR FILING DATE: 1997-09-19
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 64
/ LENGTH: 504
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-127-900A-64

Query Match      100.0%; Score 2615; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 8.1e-253;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNNLQLRPGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNNLQLRPGALYSGLMRL 60

QY 61 SKKYGPFVFTYLGPRWPVVVLVGOEAVREALGGOAEFSGRGTVAMLEGTFDGHGVFFSN 120
Db 61 SKKYGPFVFTYLGPRWPVVVLVGOEAVREALGGOAEFSGRGTVAMLEGTFDGHGVFFSN 120

QY 121 GERWROLRKFTMLALRDGLGKKEGEBELIOAEARCLVETFGTEGPRPDPSSLAAQATSN 180
Db 121 GERWROLRKFTMLALRDGLGKKEGEBELIOAEARCLVETFGTEGPRPDPSSLAAQATSN 180

QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAFTVRQVQOQHONLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
Db 241 VSTLAFTVRQVQOQHONLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300

QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKWBRELNRELGAGQAPSLGDRTRLPYTDVILHEA 360
Db 301 FAGTMTVSTTVGYTLLLMKYPHVQKWBRELNRELGAGQAPSLGDRTRLPYTDVILHEA 360

QY 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420
Db 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420

QY 421 DGRFRKHEAFLPFLSGKRVCLGELAKAELFLFTTILQAFSLESPPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAFLPFLSGKRVCLGELAKAELFLFTTILQAFSLESPPDPTLSLKPTVSG 480

QY 481 LFNIPPAFOLQVRPTDLHSTTQTR 504
Db 481 LFNIPPAFOLQVRPTDLHSTTQTR 504
```

RESULT 15

US-10-128-685A-64

Sequence 64, Application US/10128685A

Publication No. US20030203430A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tuma, Daniel

```
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P330R1C116
/ CURRENT APPLICATION NUMBER: US/10/128,685A
/ CURRENT FILING DATE: 2002-04-23
/ PRIOR APPLICATION NUMBER: 60/049911
/ PRIOR FILING DATE: 1997-06-18
/ PRIOR APPLICATION NUMBER: 60/056974
/ PRIOR FILING DATE: 1997-08-26
/ PRIOR APPLICATION NUMBER: 60/059113
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059115
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059117
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059122
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059184
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059352
/ PRIOR FILING DATE: 1997-09-19
/ PRIOR APPLICATION NUMBER: 60/059588
/ PRIOR FILING DATE: 1997-09-19
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 64
/ LENGTH: 504
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-128-685A-64
```

Query Match 100.0%; Score 2615; DB 12; Length 504;

Best Local Similarity 100.0%; Pred. No. 8.1e-253;

Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNNLQLRPGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSGTRARGHLPPTPLPLGNNLQLRPGALYSGLMRL 60

QY 61 SKKYGPFVFTYLGPRWPVVVLVGOEAVREALGGOAEFSGRGTVAMLEGTFDGHGVFFSN 120
Db 61 SKKYGPFVFTYLGPRWPVVVLVGOEAVREALGGOAEFSGRGTVAMLEGTFDGHGVFFSN 120

QY 121 GERWROLRKFTMLALRDGLGKKEGEBELIOAEARCLVETFGTEGPRPDPSSLAAQATSN 180
Db 121 GERWROLRKFTMLALRDGLGKKEGEBELIOAEARCLVETFGTEGPRPDPSSLAAQATSN 180

QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAFTVRQVQOQHONLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
Db 241 VSTLAFTVRQVQOQHONLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300

QY 301 FAGTMTVSTTVGYTLLLMKYPHVQKWBRELNRELGAGQAPSLGDRTRLPYTDVILHEA 360
Db 301 FAGTMTVSTTVGYTLLLMKYPHVQKWBRELNRELGAGQAPSLGDRTRLPYTDVILHEA 360

QY 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420
Db 361 QRLALVPMGIPRTLMTTRFRGYTLPOGTEVPFLGSLIHDNPIFKHPEEFNPDRLDA 420

QY 421 DGRFRKHEAFLPFLSGKRVCLGELAKAELFLFTTILQAFSLESPPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAFLPFLSGKRVCLGELAKAELFLFTTILQAFSLESPPDPTLSLKPTVSG 480

QY 481 LFNIPPAFOLQVRPTDLHSTTQTR 504
```

Db 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

Search completed: September 15, 2004, 09:33:07
Job time : 130 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:18:40 ; Search time 33 Seconds
(without alignments)
788.469 Million cell updates/sec

Title: US-10-669-693-2
Perfect score: 2615
Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1231.5	47.1	494	US-09-126-420A-26	Sequence 26, Appl
2	1092.5	41.8	490	US-08-201-118-13	Sequence 13, Appl
3	1092.5	41.8	490	US-08-238-821B-13	Sequence 13, Appl
4	1092.5	41.8	490	PCT-US95-05744-13	Sequence 13, Appl
5	1087.5	41.6	490	US-08-201-118-3	Sequence 3, Appl
6	1087.5	41.6	490	US-08-238-821B-3	Sequence 3, Appl
7	1087.5	41.6	490	PCT-US95-05744-3	Sequence 3, Appl
8	1085.5	41.5	490	US-08-201-118-9	Sequence 9, Appl
9	1085.5	41.5	490	US-08-238-821B-9	Sequence 9, Appl
10	1085.5	41.5	490	PCT-US95-05744-9	Sequence 9, Appl
11	1084.5	41.5	490	US-08-201-118-1	Sequence 1, Appl
12	1084.5	41.5	490	US-08-238-821B-1	Sequence 1, Appl
13	1084.5	41.5	490	PCT-US95-05744-1	Sequence 1, Appl
14	1073.5	41.1	490	US-08-201-118-5	Sequence 5, Appl
15	1073.5	41.1	490	US-08-238-821B-5	Sequence 5, Appl
16	1073.5	41.1	490	PCT-US95-05744-5	Sequence 5, Appl
17	1070.5	40.9	500	US-08-314-601-2	Sequence 2, Appl
18	1070.5	40.9	500	PCT-US95-13051-2	Sequence 2, Appl
19	1067.5	40.8	490	US-08-201-118-11	Sequence 11, Appl
20	1067.5	40.8	490	US-08-238-821B-11	Sequence 11, Appl
21	1067.5	40.8	490	PCT-US95-05744-11	Sequence 11, Appl
22	1050.5	40.2	490	US-08-201-118-7	Sequence 7, Appl
23	1050.5	40.2	490	US-08-238-821B-7	Sequence 7, Appl
24	1050.5	40.2	490	PCT-US95-05744-7	Sequence 7, Appl
25	792	30.3	504	US-09-126-420A-23	Sequence 23, Appl
26	679	26.0	502	US-09-304-615-69	Sequence 69, Appl
27	562.5	21.5	512	US-08-194-981E-5	Sequence 5, Appl

28	544.5	20.8	524	4	US-09-126-420A-24	Sequence 24, Appl
29	502	19.2	504	4	US-09-976-594-642	Sequence 642, Appl
30	439	16.8	508	4	US-09-126-420A-25	Sequence 25, Appl
31	421	16.1	501	3	US-09-158-767-15	Sequence 15, Appl
32	421	16.1	501	3	US-09-158-767-16	Sequence 16, Appl
33	421	16.1	501	3	US-09-158-767-17	Sequence 17, Appl
34	421	16.1	501	3	US-09-158-767-18	Sequence 18, Appl
35	417	15.9	504	4	US-09-499-302A-9	Sequence 9, Appl
36	405.5	15.5	506	4	US-09-672-785-2	Sequence 2, Appl
37	401.5	15.4	490	4	US-09-126-420A-16	Sequence 16, Appl
38	386.5	14.8	476	1	US-08-313-075A-30	Sequence 30, Appl
39	378	14.5	471	4	US-09-126-420A-20	Sequence 20, Appl
40	378	14.5	495	1	US-08-532-065B-2	Sequence 2, Appl
41	377	14.4	506	4	US-09-672-785-8	Sequence 8, Appl
42	374.5	14.3	505	4	US-09-627-216A-12	Sequence 12, Appl
43	374.5	14.3	505	4	US-09-765-873A-12	Sequence 12, Appl
44	374	14.3	531	4	US-09-380-420C-2	Sequence 2, Appl
45	374	14.3	531	4	US-09-899-642A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-126-420A-26
; Sequence 26, Application US/09126420A
; Patent No. 6376753
; GENERAL INFORMATION:
; APPLICANT: BATARD, YANNICK
; APPLICANT: ROBINEAU, TIRURCE
; APPLICANT: DUREST, FRANCIS
; APPLICANT: WERCK-REICHART, DANIELE
; APPLICANT: DIDIERJEAN, LUC
; TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP76B1 FROM HELIANTHUS
; TITLE OF INVENTION: TUBEROSUS AND ITS APPLICATIONS AS BIOCATALYST IN
; TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL
; TITLE OF INVENTION: POLLUTANTS AND FOR ALTERING THE RESISTANCE OF PLANTS S
; FILE REFERENCE: 03715.0032
; CURRENT APPLICATION NUMBER: US/09/126.420A
; CURRENT FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/054.351
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 26
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-126-420A-26

Query Match 47.1%; Score 1231.5; DB 4; Length 494;
Best Local Similarity 49.0%; Pred. No. 4.3e-122;
Matches 242; Conservative 93; Mismatches 150; Indels 9; Gaps 4;
QY 1 MEATGTWALLALALLLLTALSQT-----RARGHLPDGTPLPLGNLLQLRPGALYS 56
Db 1 MLAGS---LLVASVAFSLVLMVSKQKLSKLPDGTPLPLFGNYQLNTERKMYSS 57
QY 57 LMELSKYGPVFTYILGPMRPVVVLVGOEAVREALGQAEERFSGRGTVAMLEGTFDHGV 116
Db 58 LMKISORYGPVFTIHLGP-RRVVLCGQAVKALVDQAEFSGRGEQATFDLFLKGYGV 116
QY 117 FFSNGERWRLKFTTLMALDLGMKREBELTQAEARCLVETFOCTEGRPFPPDLLAQ 176
Db 117 AFSSGERAQLRFSIATLRDFGVKRGIEERLQEEAGFLIESFRKINGALIDPTFLSR 176
QY 177 ATSNVCSLLFLGRFSYDEKFEQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLGPHKQ 236
Db 177 TVSNVISSVFGDRFDYEDKEFLSLRLMMLGSGFQFATSTGTQLYEMFSSVMKHLPGPQQ 236
QY 237 LLHHVSTLAAFTVRQVQHQGNLDASGPARDLVADAFLLKMAQEEQNPGTFTNNMLMTV 296
Db 237 AFKEQGLEDFITTKVEQNRQLDNPSP-RDFIDSLFRLMEKKNPNTFYMKNLVLTT 295

```
QY 297 IYLLFAGTMTYSTTGYTLLLLMKYPHVQKWRBELNRELGAGQAPSLGDRTRLPYTDAY 356
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 296 LNLFFAGTETVSTLRYGFLLLMRHPDIEAKVHEIDRVIGRNRQAKYEDRMKMPYTEAV 355
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 357 LHEAQRLALVPMGIPRITLMTTRRGTYLPGQTEVFPFLGSLILHDPNIEKHPEEFNPDR 416
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 356 IHEIQRFADMI PMGLARRVTDKTFREFLLPKGTGVFFMLGSLVLPKDFSNPNDFNPKH 415
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 417 FLDADGRERKEAFPSISLGRKVCILGEGLAELFLFTTILQAFSLSPCPPDPLSLKP 476
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 416 FLDDKGQKKSDAFVPSIGKRYCGEGGLARMELFLFLTNIMQNECFKSPQAPQDIDVSP 475
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 477 TVSGLFNIPPAFQL 490
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 476 RLVGFAPIPPNYTM 489
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 2
US-08-201-118-13
; Sequence 13, Application US/08201118
; Patent No. 5786191
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; TITLE OF INVENTION: CLONING AND EXPRESSION OF COMPLEMENTARY
; TITLE OF INVENTION: DNAS FOR MULTIPLE MEMBERS OF THE HUMAN CYTOCHROME P450 2C
; TITLE OF INVENTION: SUBFAMILY
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/201,118
; FILING DATE: 22-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 15280-192-1
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
US-08-201-118-13
Query Match 41.8%; Score 1092.5; DB 1; Length 490;
Best Local Similarity 45.0%; Pred. No. 2.6e-107;
Matches 217; Conservative 92; Mismatches 170; Indels 3; Gaps 3;

QY 9 LLLALALLLTALSCTRAGHLPDGTPLPLGNLLQLRPGALYGLMRLSKYGVF 68
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 7 LVLCLSCLLLSLRQSS-GRGKLPDGTPLPLXIGNILQIDKDISKSLTNKXKVGPF 65
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```

; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
US-08-238-821B-13

Query Match          41.8%; Score 1092.5; DB 2; Length 490;
Best Local Similarity 45.0%; Pred. No. 2.6e-107;
Matches 217; Conservative 92; Mismatches 170; Indels 3; Gaps 3;

QY 9 LLLAALLLLTALSGTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRLSKYGVFV 68
Db 7 LVLCLSCLLLSLRQSS-GRGKLPFGPTPLPXIGNILQIDKXSKSLTNKSKYGVFV 65
QY 69 TIYLGWPRPVVLVQGAVERALGGQAEFGSGRTVAMLEGTFDGHGVFFSNGERWQLR 128
Db 66 TXYFG-LKPIVVLHGYEAVKEALIDLGEFGSGRGXFPPLAERANKXGIVFSGRWKEIR 124
QY 129 KFTMLALRDLGMKGREGELIQAEARCLVETFOGTGREGPDPSPSLLLAQATSNVVCSLIFG 188
Db 125 RFLSMTLRNFGMKRSIEDRVQEEARCLVEELRTKASPCDPTILGCPCNVICSIFH 184
QY 189 LRFSEYDEKFEQAVVRAAGTLLGVSSQGGQYEMFSWFLRPLPGPHKQLLHVSTLAAFT 248
Db 185 KRPDYKQQFNLNMEKXNENIRILSSPWIQXCNFPXXIDYFPGTHNKLKNVAFMKSII 244
QY 249 VROVOQHGNLDASGPARDLVDAFLIKMAQEQNPQGTFTNKNMLMTVIYLLFAGTMTVS 308
Db 245 LEKVEHQESXDMNRP-RDFIDCLIKMEKXNQSEFTIESLXXTXDLFGAGTETS 303
QY 309 TTVGYTLLLLMKYPHVQKWRVRELNRGLGAGQAPSLGDRTRPLPYTDAVLHEAQRLLALVP 368
Db 304 TTLRYLXLLLLKHPEVTAKEVIERVIGNRSPCMQDRSHMPTDVAHVEXQRYIDLCP 363
QY 369 MGIPRTLMRTTRRGYTLPGTGVFPLLSILHDPNIFKHPEEFNPDRLDADGRFRKHE 428
Db 364 TSLPHAVTCDVKFRNYLIPKGTITLTSVLHDXKEFPNEMFDPGHFLDXGKNFKSD 423
QY 429 AFLPFLSGKVCGLGEGLAKAELEFPTTILQAFSLSPCPDPTLSLXPTVSGLNIPPAF 488
Db 424 YFMPFSGAKRICVGEGLARMELFLFTTILQNFNLKSLVDPKXLDTPPVVNGFASVPFFY 483
QY 489 QL 490
Db 484 QL 485

```

```

RESULT 4
PCT-US95-05744-13
; Sequence 13, Application PC/TUS9505744
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; APPLICANT: DE MORRIS, Sonia M.F.
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Hourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05744
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,821
; FILING DATE: 06-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/201,118
; FILING DATE: 22-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 15280-192-1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
PCT-US95-05744-13

```

```

Query Match          41.8%; Score 1092.5; DB 5; Length 490;
Best Local Similarity 45.0%; Pred. No. 2.6e-107;
Matches 217; Conservative 92; Mismatches 170; Indels 3; Gaps 3;

QY 9 LLLAALLLLTALSGTRARGHLPPGPTPLPLGNLLQLRPGALYSGLMRLSKYGVFV 68
Db 7 LVLCLSCLLLSLRQSS-GRGKLPFGPTPLPXIGNILQIDKXSKSLTNKSKYGVFV 65
QY 69 TIYLGWPRPVVLVQGAVERALGGQAEFGSGRTVAMLEGTFDGHGVFFSNGERWQLR 128
Db 66 TXYFG-LKPIVVLHGYEAVKEALIDLGEFGSGRGXFPPLAERANKXGIVFSGRWKEIR 124
QY 129 KFTMLALRDLGMKGREGELIQAEARCLVETFOGTGREGPDPSPSLLLAQATSNVVCSLIFG 188
Db 125 RFLSMTLRNFGMKRSIEDRVQEEARCLVEELRTKASPCDPTILGCPCNVICSIFH 184
QY 189 LRFSEYDEKFEQAVVRAAGTLLGVSSQGGQYEMFSWFLRPLPGPHKQLLHVSTLAAFT 248
Db 185 KRPDYKQQFNLNMEKXNENIRILSSPWIQXCNFPXXIDYFPGTHNKLKNVAFMKSII 244
QY 249 VROVOQHGNLDASGPARDLVDAFLIKMAQEQNPQGTFTNKNMLMTVIYLLFAGTMTVS 308
Db 245 LEKVEHQESXDMNRP-RDFIDCLIKMEKXNQSEFTIESLXXTXDLFGAGTETS 303
QY 309 TTVGYTLLLLMKYPHVQKWRVRELNRGLGAGQAPSLGDRTRPLPYTDAVLHEAQRLLALVP 368
Db 304 TTLRYLXLLLLKHPEVTAKEVIERVIGNRSPCMQDRSHMPTDVAHVEXQRYIDLCP 363
QY 369 MGIPRTLMRTTRRGYTLPGTGVFPLLSILHDPNIFKHPEEFNPDRLDADGRFRKHE 428
Db 364 TSLPHAVTCDVKFRNYLIPKGTITLTSVLHDXKEFPNEMFDPGHFLDXGKNFKSD 423
QY 429 AFLPFLSGKVCGLGEGLAKAELEFPTTILQAFSLSPCPDPTLSLXPTVSGLNIPPAF 488
Db 424 YFMPFSGAKRICVGEGLARMELFLFTTILQNFNLKSLVDPKXLDTPPVVNGFASVPFFY 483
QY 489 QL 490
Db 484 QL 485

```

Db 304 TTLVALLLLKHPVTAQVQEEIERTVGNRNSPCMQDRSHMPYTDVAVHVEVQRYLDLLP 363
Qy 369 MGIPRTLMRTFRFGYTLPOGTEVFFPLGLSGTLHDPNIFKHPBEFNDRFLDADGRFKHE 428
Db 364 TSLPHAVTCDIKFRNYLIPKGTTLILSTSVLHDKFEPNEMFDPHFLDEGNGFKSK 423
Qy 429 AFLPFSGLKRVCLGEGLAKELEFFFTTILQAFSLSPCPDPTLSLKPTVSGLFNIPPAF 488
Db 424 YFMFSAKRICVGEALAGMELFLFLTSILQNFNLKSLVDPKDLDTTPVANGFASVPFFY 483
Qy 489 QL 490
Db 484 QL 485
RESULT 6
US-08-238-821B-3
; Sequence 3, Application US/0823821B
; Patent No. 5912120
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-
; TITLE OF INVENTION: MEPHENYTOIN METABOLISM
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US/08/238,821B
; APPLICATION NUMBER: US 08/201,118
; FILING DATE: 22-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 15280-192110US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-238-821B-3
Query Match 41.6%; Score 1087.5; DB 2; Length 490;
Best Local Similarity 44.8%; Pred. No. 8.9e-107;
Matches 216; Conservative 99; Mismatches 164; Indels 3; Gaps 3;
Qy 9 LLLALLLLLTALSGTRAGHLPGGTPPLPILGNLLQLRPGALYSGLMRLSKYGPVF 68
Db 7 LVLCLSCLLLSLWRQSS-GRGKLPGGTLPVIGNILQIGIKDLSKLTNLSKVYGPVF 65

Db 304 TTLVALLLLKHPVTAQVQEEIERTVGNRNSPCMQDRSHMPYTDVAVHVEVQRYLDLLP 363
Qy 369 MGIPRTLMRTFRFGYTLPOGTEVFFPLGLSGTLHDPNIFKHPBEFNDRFLDADGRFKHE 428
Db 364 TSLPHAVTCDIKFRNYLIPKGTTLILSTSVLHDKFEPNEMFDPHFLDEGNGFKSK 423
Qy 429 AFLPFSGLKRVCLGEGLAKELEFFFTTILQAFSLSPCPDPTLSLKPTVSGLFNIPPAF 488
Db 424 YFMFSAKRICVGEALAGMELFLFLTSILQNFNLKSLVDPKDLDTTPVANGFASVPFFY 483
Qy 489 QL 490
Db 484 QL 485
RESULT 5
US-08-201-118-3
; Sequence 3, Application US/08201118
; Patent No. 5786191
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; TITLE OF INVENTION: CLONING AND EXPRESSION OF COMPLEMENTARY
; TITLE OF INVENTION: DNAS FOR MULTIPLE MEMBERS OF THE HUMAN CYTOCHROME P450 2C
; TITLE OF INVENTION: SUBFAMILY
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 22-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/864,962
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 15280-192-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-201-118-3
Query Match 41.6%; Score 1087.5; DB 1; Length 490;
Best Local Similarity 44.8%; Pred. No. 8.9e-107; Mismatches 164; Indels 3; Gaps 3;
Matches 216; Conservative 99; Mismatches 164; Indels 3; Gaps 3;
Qy 9 LLLALLLLLTALSGTRAGHLPGGTPPLPILGNLLQLRPGALYSGLMRLSKYGPVF 68
Db 7 LVLCLSCLLLSLWRQSS-GRGKLPGGTLPVIGNILQIGIKDLSKLTNLSKVYGPVF 65
Qy 69 TIYGLPWRPVVLVQGVAREALGGQAEPSGRTVMLEGTDFGCHGVFSPSNERWQLR 128
Db 66 TLYFG-LKPIVLVHGYEAVKEALIDLGEFSGRGIPLAERANRGFIVFSNGKKWKEIR 124
Qy 129 KFTMLALRDLMGKREGFEELIQAEARCLVETFGTEGPPDPSLLLAQATSNVVCSLFFG 188
Db 125 RFSIMTLRNGMGKRSIEDRVQEARCLVBEELKTKASPCDPTFILGCAPCNVICSIFH 184
Qy 189 LRFSEYDEKFAVVAAGGTLGVSSQGGQTYEMFSLRPLPGPHKQLLHHVSTLAFT 248
Db 185 KRFYKQDQFLNLMKLNENIKILSPWICNNFSPDIIDYFPGTHNKLKLVAFMKSYI 244
Qy 249 VRQVQHQGNLDASGPARDLVDAFLKWAQEQNPCTEFTKNMLMTVIVLLFAGTWTVS 308
Db 245 LEKVKEQESMDMNNP-QDFIDCFMLMKVEKHKHQPSEFTIESLENTAVDLFGAGTETS 303
Qy 309 TTGVYTLTLLMKYPHVQKWRVEELNRELGAQAPSLGDRTRLPYTDVAVLHEAQRLALVP 369


```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/201,118
; FILING DATE: 22-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 15280-192-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-201-118-9

Query Match 41.5%; Score 1085.5; DB 1; Length 490;
Best Local Similarity 44.6%; Pred. No. 1.5e-106;
Matches 215; Conservative 100; Mismatches 164; Indels 3; Gaps 3;

QY 9 LLLALALLLLTALSGRARHLPPTPLLLGNLLQLRPGALYSGLMRLSKYGPVF 68
DB 7 LVLCSCLLLSLWRQSS-GRGKLPPTPLPVIGNILQIGIKDISKSLTNLSKVYGPVF 65
QY 69 TTYLGPWRPVVVLVGOEAVREALGQAEFFSGRGTVAMLEGTFDGHGVFSGNWRQLR 128
DB 66 TLVYFG-LKPIVVLHGVEAVKEALIDLGEFFSGRGIFPLAERANRGFGIVFSGKWKKEIR 124
QY 129 KFTMLALRDLGMKRGEGELIQAEARCLVETFGTEGRPDPSSLQAQATSNVVCLLFG 188
DB 125 RFSLMTLRNFGMKRSIEDRVQEARCLVEELRKTASPCDPTFILGCAPCNVICSIIFH 184
QY 189 LRFSEYDEKFOAVVRAAGGTLGVSSGGQGYEMFSWFLRPLGPHKQLLHVSTLAAFT 248
DB 185 KRFYDKQQOQFLNLMKLNENIKILSSPWIQICNNFSPIIDYPPGTHNKLKNVAFMKSYI 244
QY 249 VRQVQHQGNLDASGPARDLVDAFLKMAQEQNPGTEFTNKNMLTVIYLLFAGTMTVS 308
DB 245 LEKVKEHQESMDMNNP-QDFIDCFLMKMEKEKHQSEFTIESLENTAVDLFGAGTETTS 303
QY 309 TTVGYTLLLLMKYPHYQKWRBELNRELGAQAPSLGDRTRLPYTDVAVLHEAQRLLALVP 368
DB 304 TTLRYALLLLKHPEVTAKVQEEIERVIGNRSPCMQDRSHMPYTDVAVHEVQRYIDLDP 363
QY 369 MGIPRLMTTRRGYTLQCGTEVFLGLSILHDPNIEKHPENFDRFLDADGRFRKHE 428
DB 364 TSLPHAVTCDIKPRNVLIPKGTILLISLTVLHDNKEFFPENPFDHHLDECGNFKSK 423
QY 429 AFLPFLGKRVCLGEGAKAELEFFTTILQAFSLSPCPDPTLSLKTPTVSGLFNIPPAF 488
DB 424 YWPFPSAGKRCVGEALAGNELFLTSLIQNENLKSVDLPKNDLITPPVNGFASVPFFY 483
QY 489 QL 490
DB 484 QL 485

RESULT 9
US-08-238-821B-9
; Sequence 9, Application US/08238821B
; Patent No. 5912120
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; APPLICANT: DE MORAIS, Sonia M.F.
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-
; TITLE OF INVENTION: MEPHENYTOIN METABOLISM
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,821B
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/201,118
; FILING DATE: 22-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 15280-192110US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-238-821B-9

Query Match 41.5%; Score 1085.5; DB 2; Length 490;
Best Local Similarity 44.6%; Pred. No. 1.5e-106;
Matches 215; Conservative 100; Mismatches 164; Indels 3; Gaps 3;

QY 9 LLLALALLLLTALSGRARHLPPTPLLLGNLLQLRPGALYSGLMRLSKYGPVF 68
DB 7 LVLCSCLLLSLWRQSS-GRGKLPPTPLPVIGNILQIGIKDISKSLTNLSKVYGPVF 65
QY 69 TTYLGPWRPVVVLVGOEAVREALGQAEFFSGRGTVAMLEGTFDGHGVFSGNWRQLR 128
DB 66 TLVYFG-LKPIVVLHGVEAVKEALIDLGEFFSGRGIFPLAERANRGFGIVFSGKWKKEIR 124
QY 129 KFTMLALRDLGMKRGEGELIQAEARCLVETFGTEGRPDPSSLQAQATSNVVCLLFG 188
DB 125 RFSLMTLRNFGMKRSIEDRVQEARCLVEELRKTASPCDPTFILGCAPCNVICSIIFH 184
QY 189 LRFSEYDEKFOAVVRAAGGTLGVSSGGQGYEMFSWFLRPLGPHKQLLHVSTLAAFT 248
DB 185 KRFYDKQQOQFLNLMKLNENIKILSSPWIQICNNFSPIIDYPPGTHNKLKNVAFMKSYI 244
QY 249 VRQVQHQGNLDASGPARDLVDAFLKMAQEQNPGTEFTNKNMLTVIYLLFAGTMTVS 308
DB 245 LEKVKEHQESMDMNNP-QDFIDCFLMKMEKEKHQSEFTIESLENTAVDLFGAGTETTS 303
QY 309 TTVGYTLLLLMKYPHYQKWRBELNRELGAQAPSLGDRTRLPYTDVAVLHEAQRLLALVP 368
DB 304 TTLRYALLLLKHPEVTAKVQEEIERVIGNRSPCMQDRSHMPYTDVAVHEVQRYIDLDP 363
```


LENGTH: 490 amino acids	41.5%;	Score 1084.5;	DB 1;	Length 490;
TYPE: amino acid	Best Local Similarity	45.0%;	Pred. No. 1.9e-106;	
TOPOLOGY: unknown	Matches 217;	Conservative	96;	Mismatches 166;
MOLECULE TYPE: protein				Indels 3; Gaps 3;
US-08-201-118-1				
Query Match				
Best Local Similarity				
Matches 217;				
Conservative				
Indels 3;				
Gaps 3;				
9 LLLAALALLLTALSGTRAGHLPPOPTPLPGLNLLQLRPGALYSGLMRLSKKYPVF 68				
7 LVLCISCLLLSIWRQSS-GRGKLPPGPTPLPVGNILQIDIKDVSKSLTNLSKIYGPVF 65				
69 TIYLGPRPVPVVLVGQAVREALGQAEFEFGSGRTVAMLEGTDFGHGVFFSNGERWRQLR 128				
66 TIYGLER-MVVLHGVEYVKEALIDLGEFEFGSGRHFFLAERANRGFIVFSGKRWKEIR 124				
129 KFTMLALRDLGMGKREGEELIQAERCLVETFOGTEGRPFDPSSLILAQATSNVCSLLFG 188				
125 RFLMTLRFNGMGKRSIEDRVQEEARCLVEELRKTASPCDPTFILGCAPCNVICSIIFQ 184				
189 LRFSEYDEKFOAVRAAGTLLGVSSGGQTYEMFSWFLRPLPGPHKQLLHVSTLAAFT 248				
185 KRFYDKQDFNLMEKLNENRIVSTPWICNNFTIIDYFFGTHNKLKLNLAFAVESDI 244				
249 VRQVOHQGNLDASGPARDLVDAFLKMAQEEQNPQTEFTNKNMLMTVIYLLFAGTWTVS 308				
245 LEKVKHEQESMDINN-RDFIDCFLIKMEKEKQOQSEFTIENIVTAADLLGAGTETTS 303				
309 TTVGTYLILLMKYPHVQKWVREELNRELGAQASLGDRTLPYTDVLAHEAQRLLALVP 368				
304 TTLRYALLLLKHPEVTAKVQEEIERVIGNRNSPCMDQGHMPTIDAVHEVQRYIDLIP 363				
369 MGIPRTLMTTRFRGYTLPOGTEVPPLIGSLIHPDNIFKHPEEFNPDORFLDADGRFKHE 428				
364 TSLPHAVTCDVKFRNYLIPKGTITLTSVLHNDKKEFPNDEMDFRHLDEGNGFKSN 423				
429 AFLPSLGRKVCGLGKAELELFFFTIILQAFSLSPCPDPTLSLXPTVSGLFNIPPAF 488				
424 YMPFSGAKRIICVGEGLARMELFLTLFQNFNLKSLIDPKDLDTTPVVGNGFASVPFFY 483				
489 QL 490				
484 QL 485				
RESULT 12				
US-08-238-821B-1				
Sequence 1. Application US/08238821B				
Patent No. 5912120				
GENERAL INFORMATION:				
APPLICANT: GOLDSTEIN, Joyce A.				
APPLICANT: ROMKES-SPARKS, Marjorie				
APPLICANT: DE MORAIS, Sonia M.F.				
TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN				
TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-				
TITLE OF INVENTION: MEPHENYTOIN METABOLISM				
NUMBER OF SEQUENCES: 61				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: Townsend and Crew LLP				
STREET: Two Embarcadero Center, 8th Floor				
CITY: San Francisco				
STATE: California				
COUNTRY: US				
ZIP: 94111				
COMPUTER READABLE FORM: disk				
MEDIUM TYPE: Floppy disk				
COMPUTER: IBM PC compatible				
OPERATING SYSTEM: PC-DOS/MS-DOS				
SOFTWARE: Patent In Release #1.0, Version #1.25				
CURRENT APPLICATION DATA:				
INVENTION NUMBER: US/08/238 821B				

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 09:14:35 ; Search time 128 Seconds

(without alignments)
1112.531 Million cell updates/sec

Title: US-10-669-693-2

Perfect score: 2615

Sequence: 1 MEATGTWALLALALLLLT.....PPAFQLQVRPTDLHSTTQTR 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2615	100.0	504	4 AAU12203	Aau12203 Human PRO
2	2615	100.0	504	5 AAE28282	Aae28282 Human PRO
3	2615	100.0	504	5 AAE15439	Aae15439 Human PRO
4	2615	100.0	504	6 ABO17647	Abol17647 Novel hum
5	2615	100.0	504	6 ABU10222	Abu10222 Human cyt
6	2615	100.0	504	6 ABU80901	Abu80901 Human PRO
7	2615	100.0	504	6 ABU66601	Abu66601 Human PRO
8	2615	100.0	504	6 ABU59682	Abu59682 Novel sec
9	2615	100.0	504	6 ABU24872	Abu24872 Human sec
10	2615	100.0	504	6 ABU66877	Abu66877 Human sec
11	2615	100.0	504	6 ADA45583	Ada45583 Novel hum
12	2615	100.0	504	6 ADA76014	Ada76014 Human PRO
13	2615	100.0	504	6 ADA18664	Ada18664 Human PRO
14	2615	100.0	504	6 ADA61287	Ada61287 Homo sapi
15	2615	100.0	504	6 ADA19072	Ada19072 Novel hum
16	2615	100.0	504	6 ADB27613	Adb27613 Human PRO
17	2615	100.0	504	6 ADB86092	Adb86092 Novel hum
18	2615	100.0	504	6 ADB15656	Adb15656 Human PRO
19	2615	100.0	504	6 ADA47442	Ada47442 Human PRO
20	2615	100.0	504	6 ADA67237	Ada67237 Human PRO
21	2615	100.0	504	6 ADB30244	Adb30244 Human PRO
22	2615	100.0	504	6 ADA85540	Ada85540 Novel hum
23	2615	100.0	504	6 ADA96752	Ada96752 Human PRO
24	2615	100.0	504	6 ADA79056	Ada79056 Human PRO
25	2615	100.0	504	6 ADA87195	Ada87195 Novel hum

ALIGNMENTS

RESULT 1

AAU12203

ID AAU12203 standard; protein; 504 AA.

AC AAU12203;

XX 24-OCT-2001 (first entry)

DT Human PRO1906 polypeptide sequence.

DE Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;

KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;

KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;

KW A-peptide; factor VIIA; gene therapy.

XX Homo sapiens.

XX WO200140466-A2.

PN 07-JUN-2001.

XX 01-DEC-2000; 2000WO-US032678.

XX 01-DEC-1999; 99WO-US028301.

PR 01-DEC-1999; 99WO-US028634.

PR 02-DEC-1999; 99WO-US028551.

PR 02-DEC-1999; 99WO-US028564.

PR 02-DEC-1999; 99WO-US028565.

PR 09-DEC-1999; 99US-0170262P.

PR 16-DEC-1999; 99WO-US030095.

PR 20-DEC-1999; 99WO-US030911.

PR 30-DEC-1999; 99WO-US031243.

PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000277.

PR 11-FEB-2000; 2000WO-US000376.

PR 18-FEB-2000; 2000WO-US003565.

PR 18-FEB-2000; 2000WO-US004341.

PR 22-FEB-2000; 2000WO-US004342.

PR 24-FEB-2000; 2000WO-US004914.

PR 24-FEB-2000; 2000WO-US005004.

PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005841.

PR 03-MAR-2000; 2000US-0187202P.

PR 10-MAR-2000; 2000WO-US006319.

PR 15-MAR-2000; 2000WO-US006884.

ADb16397 Human PRO
Ada91489 Novel hum
ADb14552 Human PRO
ADb18513 Novel hum
Ada93728 Human PRO
ADb19624 Novel hum
ADb12936 Human PRO
Abo43180 Novel hum
Ada74130 Human PRO
ADb24423 Human PRO
Ada81947 Human PRO
Ada74910 Human PRO
Ada84988 Novel hum
Ada84436 Novel hum
ADb29692 Human PRO
Ada80220 Human PRO
Ada75462 Human PRO
Ada46687 Human PRO
ADb24983 Human PRO
Ada93159 Human PRO

PR	20-MAR-2000;	2000WO-US007377.	
PR	21-MAR-2000;	2000WO-US007532.	
PR	30-MAR-2000;	2000WO-US008439.	
PR	17-MAY-2000;	2000WO-US013705.	
PR	22-MAY-2000;	2000WO-US014042.	
PR	30-MAY-2000;	2000WO-US014941.	
PR	02-JUN-2000;	2000WO-US015264.	
PR	05-JUN-2000;	2000US-0209832P.	
PR	28-JUL-2000;	2000WO-US020710.	
PR	11-AUG-2000;	2000WO-US022031.	
PR	23-AUG-2000;	2000WO-US023522.	
PR	24-AUG-2000;	2000WO-US023328.	
PR	08-NOV-2000;	2000WO-US030952.	
PR	10-NOV-2000;	2000WO-US030873.	
XX		(GETH) GENENTECH INC.	
PA			
XX	Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;		
PI	Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;		
PI	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;		
XX			
DR	WPI; 2001-408281/43.		
DR	N-PSDB; AAS21275.		
XX			
XX	Isolated , secretory and transmembrane PRO polypeptide used to detect		
PT	other PRO polypeptides, link bioactive molecules to cells expressing PRO		
PT	polypeptides, and detect the presence of mammalian tumors e.g. lung,		
PT	breast, prostate, cervical.		
XX			
PS	Claim 12; Fig 64; 813pp; English.		
XX			
CC	AAU12172-AAU12446 represent novel human secretory and transmembrane PRO		
CC	polypeptides. The PRO polypeptides are useful to detect other PRO		
CC	polypeptides, to link bioactive molecules to cells expressing PRO		
CC	polypeptides, to modulate biological activities of cells expressing PRO		
CC	polypeptides, and to detect the presence of mammalian lung, colon,		
CC	breast, prostate, rectal, cervical or liver tumours by comparing PRO		
CC	polypeptide expression in a cell sample to that in a control sample. Some		
CC	of the 275 sequences are also useful to stimulate the release of tumour		
CC	necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or		
CC	differentiation of chondrocytes, the proliferation or gene expression in		
CC	pericyte cells, the release of proteoglycans from cartilage, the		
CC	proliferation of inner ear utricular supporting cells or of T-		
CC	lymphocytes, the release of a cytokine from peripheral blood monocytes		
CC	(PBMCs), or the proliferation of endothelial cells. Some of the PRO		
CC	polypeptides may modulate glucose or free fatty acid uptake by skeletal		
CC	muscle cells or by adipocytes; or inhibit binding of A-peptide to factor		
CC	VIIA. The PRO polypeptides can be used in assays to identify molecules		
CC	involved in binding interactions. The polynucleotides encoding PRO		
CC	polypeptides can be used to generate probes, antisense RNA/DNA,		
CC	transgenic or knock out animals and can be used in gene therapy		
XX			
SQ	Sequence 504 AA;		
	Query Match 100.0%; Score 2615; DB 4; Length 504;		
	Best Local Similarity 100.0%; Pred. No. 1.8e-258;		
	Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MEATGTWALLLALLLTLALSGTRAGHLPPGPTPLLLGNLLQRLPGALYGLMRL 60		
DB	1 MEATGTWALLLALLLTLALSGTRAGHLPPGPTPLLLGNLLQRLPGALYGLMRL 60		
QY	61 SKKYGVPFTIYLGWPWPVVVLVGQAEVREALGGQAEFFSGRTVAMLEGTPDHGVFFSN 120		
DB	61 SKKYGVPFTIYLGWPWPVVVLVGQAEVREALGGQAEFFSGRTVAMLEGTPDHGVFFSN 120		
QY	121 GERWRQLRFTMLALRDLCMKRGEELIQAEARCLVETFGTEGRPPDPSLLLAQATSN 180		
DB	121 GERWRQLRFTMLALRDLCMKRGEELIQAEARCLVETFGTEGRPPDPSLLLAQATSN 180		
QY	181 VVCSLLFLGRFSYEDKEFQAVVRAAGTLLGVSSGGGTVMFSLRPLFCPHKQLLHH 240		
DB	181 VVCSLLFLGRFSYEDKEFQAVVRAAGTLLGVSSGGGTVMFSLRPLFCPHKQLLHH 240		

QY	241 VSTLAFTVRQVQOQGNLDASGPARDLVDAFLKMAQEENQPGTEFTNKMLMTVIYLL 300			
DB	241 VSTLAFTVRQVQOQGNLDASGPARDLVDAFLKMAQEENQPGTEFTNKMLMTVIYLL 300			
QY	301 FAGTMTVSTTVGYTLLLLMKYPHVQKWRRELNRELGAGQAPSLGDRTRLPYTDVLAHEA 360			
DB	301 FAGTMTVSTTVGYTLLLLMKYPHVQKWRRELNRELGAGQAPSLGDRTRLPYTDVLAHEA 360			
QY	361 QRLIALVPMGIPRILMTRTTRRGYTLPGQTEVFPPLGSLILHDPNIFKHPEEFPDFLDA 420			
DB	361 QRLIALVPMGIPRILMTRTTRRGYTLPGQTEVFPPLGSLILHDPNIFKHPEEFPDFLDA 420			
QY	421 DGRFRKHEAFPLPFSGLKRVCLGEGAKAELEFFFTILQAFSLESPPDPLSLKPTVSG 480			
DB	421 DGRFRKHEAFPLPFSGLKRVCLGEGAKAELEFFFTILQAFSLESPPDPLSLKPTVSG 480			
QY	481 LFNIPPAFQLQVRPTDLHSTTQTR 504			
DB	481 LFNIPPAFQLQVRPTDLHSTTQTR 504			
	RESULT 2			
	AAE28282			
ID	AAE28282 standard; protein; 504 AA.			
XX	AAE28282;			
XX	27-DEC-2002 (first entry)			
XX	Human drug-metabolising enzyme.			
XX	Human; drug-metabolising enzyme; cytochrome p450; therapy; cytostatic;			
KW	Parkinson's disease; myotonic dystrophy; developmental defect; enzyme;			
KW	cancer; chromosome 19.			
OS	Homo sapiens.			
XX				
Key	Location/Qualifiers			
Region	5..25	/note= "Helix 1"		
Region	32..53	/note= "Leucine zipper pattern"		
Region	39..60	/note= "Leucine zipper pattern"		
Modified-site	51..56	/note= "N-myristoylation site"		
Modified-site	61..63	/note= "Protein kinase C phosphorylation site"		
Region	64..84	/note= "Helix 2"		
Modified-site	99..101	/note= "Protein kinase C phosphorylation site"		
Modified-site	109..114	/note= "N-myristoylation site"		
Modified-site	115..120	/note= "N-myristoylation site"		
Modified-site	119..122	/note= "Casein kinase II phosphorylation site"		
Modified-site	128..131	/note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"		
Modified-site	140..143	/note= "Amidation site"		
Region	170..190	/note= "Helix 3"		
Modified-site	188..193	/note= "N-myristoylation site"		
Modified-site	192..195	/note= "Casein kinase II phosphorylation site"		
Modified-site	207..212	/note= "N-myristoylation site"		
Modified-site	248..250	/note= "N-myristoylation site"		

Db 421 DGRFKHEAFPLPSLGRKVCLEGLAKAELFLFTTILQAFSLSPCPDPTLSLKPTVSG 480
Qy 481 LFNIPPAFQLOVRPTDLHSTTQTR 504
Db 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

RESULT 5
ABU10222
ID ABU10222 standard; protein; 504 AA.
XX AC ABU10222;
XX DT 28-JUL-2003 (first entry)
XX DE Human cytochrome P450, 33303.
XX KW Human; cytochrome; P450; cellular proliferation; cancer; migraine;
KW cellular differentiation; cell adhesion; cell motility; cell migration;
KW inflammation; autoimmune disorder; rheumatoid arthritis; diabetes;
KW multiple sclerosis; systemic lupus erythematosus; conjunctivitis; asthma;
KW allergy; Grave's disease; graft-versus-host disease; vascular disorder;
KW central nervous system disorder; Alzheimer's disease; anxiety disorder;
KW amnesia; depression; Parkinson's disease; obesity; cachexia; anorexia;
KW cell-signaling-associated disorder; metabolism-associated disorder;
KW steroid-associated disorder; fatty acid-associated disorder.
XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT Peptide 1..29
FT /label= Signal_sequence
FT Protein 30..504
FT /label= Mature_33303
FT Region 32..53
FT /label= Leucine_zipper
FT Domain 33..493
FT /label= Cytochrome_P450_domain
FT Modified-site 51..56
FT /note= "N-myristoylation site"
FT Modified-site 61..63
FT /note= "Protein kinase C phosphorylation site"
FT Binding-site 99..102
FT /label= Glycosaminoglycan_attachment_site
FT Modified-site 99..101
FT /note= "Protein kinase C phosphorylation site"
FT Modified-site 109..114
FT /note= "N-myristoylation site"
FT Modified-site 115..120
FT /note= "N-myristoylation site"
FT Modified-site 119..122
FT /note= "Casein kinase II phosphorylation site"
FT Modified-site 128..131
FT /note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT Modified-site 140..143
FT /note= "Amidation site"
FT Modified-site 188..193
FT /note= "N-myristoylation site"
FT Modified-site 192..195
FT /note= "Casein kinase II phosphorylation site"
FT Modified-site 207..212
FT /note= "N-myristoylation site"
FT Modified-site 248..250
FT /note= "Protein kinase C phosphorylation site"
FT Modified-site 257..261
FT /note= "N-myristoylation site"
FT Modified-site 284..289
FT /note= "N-myristoylation site"
FT Modified-site 288..290
FT /note= "Protein kinase C phosphorylation site"
FT Modified-site 339..344

FT Modified-site /note= "N-myristoylation site"
FT 343..346
FT /note= "Casein kinase II phosphorylation site"
FT Modified-site /note= "N-myristoylation site"
FT 370..375
FT Modified-site /note= "Protein kinase C phosphorylation site"
FT 378..380
FT Region 433..442
FT /note= "Cytochrome P450 cysteine haem-iron ligand signature"
FT Modified-site 435..438
FT /note= "Amidation site"
FT Modified-site 444..449
FT /note= "N-myristoylation site"
FT Modified-site 473..475
FT /note= "Protein kinase C phosphorylation site"
XX US2003022334-A1.
XX PD 30-JAN-2003.
XX XX 04-FEB-2002; 2002US-00067668.
XX PF 02-FEB-2001; 2001US-0266140P.
XX PR (GLUC/) GLUCKSMANN M A.
XX PI Glucksmann MA;
XX DR WPI; 2003-447427/42.
XX DR N-PSDB; ACA61907.
XX PT New 33312, 33303 or 32579 nucleic acid molecule encoding cytochrome P450
PT polypeptides, useful for diagnosing and treating cancer, inflammation,
PT vascular disorders, CNS disorders or metabolic disorders, and in
PT pharmacogenomics.
XX Claim 5; Page 48; 76pp; English.
XX PS The invention relates to an isolated nucleic acid molecule comprising any
XX of the 6 nucleic acid sequences, given in the specification which encodes
CC a polypeptide having any of the 3 amino acid sequences, given in the
CC specification. The nucleic acid molecule, protein and antibody are useful
CC in screening assays and predictive medicine, such as in diagnostic
CC assays, prognostic assays, monitoring clinical trials and
CC pharmacogenetics. The composition and method are used in diagnosing and
CC treating 33312, 33303 or 32579 cytochrome P450-related disorders, such as
CC cellular proliferative and/or differentiative disorders (e.g. cancer),
CC disorders associated with undesirable or deficient cell adhesion,
CC motility or migration, inflammatory or autoimmune disorders (e.g.
CC rheumatoid arthritis, diabetes, multiple sclerosis, systemic lupus
CC erythematosus, conjunctivitis, allergies, asthma, Grave's disease or
CC graft-versus-host disease), vascular disorders, central nervous system
CC disorders (e.g. Alzheimer's disease, anxiety disorders, amnesia,
CC depression, migraine or Parkinson's disease), cell-signaling-associated
CC disorders, metabolism-associated disorders (e.g. obesity, cachexia or
CC anorexia), steroid-associated disorders and fatty acid-associated
CC disorders. The present sequence represents the amino acid sequence of
CC human cytochrome P450, 33303
XX SQ Sequence 504 AA;
SQ Query Match 100.0%; Score 2615; DB 6; Length 504;
SQ Best Local Similarity 100.0%; Pred. No. 1.8e-258;
SQ Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEATGTWALLLALLLLLLLALSGTARGHLPGPTPLPLGNLLQLRPGALYSGLMRL 60
Db 1 MEATGTWALLLALLLLLLLALSGTARGHLPGPTPLPLGNLLQLRPGALYSGLMRL 60
Qy 61 SKKYGVPFTLYLGPWRPVVVLVGQEAVERALGGQAEFFSGRGTVAMLEGTGDGHVFPFN 120
Db 61 SKKYGVPFTLYLGPWRPVVVLVGQEAVERALGGQAEFFSGRGTVAMLEGTGDGHVFPFN 120

PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
XX
PA (GETH) GENENTECH INC.
XX
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2003-352836/33.
DR N-PEDB; ACA67025.
XX
XX New isolated PRO polypeptide useful for treating diabetes, rheumatoid
PT arthritis, sports injuries, obesity, hearing loss in mammals, stroke, or
PT heart attack.
XX
XX Claim 12; Fig 64; 643pp; English.
XX
XX The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The PRO
CC polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides and polynucleotides are useful for preparing a medicament
CC useful in the treatment of diabetes, bone and/or cartilage disorders
CC (e.g. rheumatoid arthritis, sports injuries, osteoarthritis), obesity,
CC hyper- or hypo-insulinaemia, hearing loss, and coagulation disorders
CC assays for PRO, by detecting its expression in specific cells, tissues or
CC serum, and for affinity purification of PRO from recombinant cell culture
CC or natural sources. ABU8670-ABU81144 represent the human PRO
CC polypeptides of the invention. Note: The sequence data for this patent
CC was obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/psipsdIDEntry.html
XX
XX Sequence 504 AA;
SQ
Query Match 100.0%; Score 2615; DB 6; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.8e-256;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEATGTWALLALALLLLTALSGLTRARGHLPPTPLPLGMLQLRPGALYGLMRL 60
DB 1 MEATGTWALLALALLLLTALSGLTRARGHLPPTPLPLGMLQLRPGALYGLMRL 60
QY 61 SKKYGVFTIYLGFWRPVVVLVQEA VREALGQAEFEFSRGRTVAMLEGTFDGHGVFFSN 120
DB 61 SKKYGVFTIYLGFWRPVVVLVQEA VREALGQAEFEFSRGRTVAMLEGTFDGHGVFFSN 120
QY 121 GERWRQLRKFTMLALRDGLMGKGECELIQAEARCLVETFGTEGPEFDPDSLLLAQATSN 180
DB 121 GERWRQLRKFTMLALRDGLMGKGECELIQAEARCLVETFGTEGPEFDPDSLLLAQATSN 180
QY 181 VVCSLLFGLRFSYEDKEFQAVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRFSYEDKEFQAVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300
DB 241 VSTLAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEQNPGTEFTNKMLMTVIYLL 300
QY 301 FAGTMTVSTTVGYTLLLMKYBHVQKWVREELNRELGAQAPSLGDRTRLPYTDVALLHEA 360
DB 301 FAGTMTVSTTVGYTLLLMKYBHVQKWVREELNRELGAQAPSLGDRTRLPYTDVALLHEA 360
QY 361 QRLALVPMGIPRTIMRTTRFRGYTLPGQTEVFFLLGSILHDPNIFKHPEEFPNDRFLDA 420

Db 361 QRLALVPMGIPRTIMRTTRFRGYTLPGQTEVFFLLGSILHDPNIFKHPEEFPNDRFLDA 420
QY 421 DGRERKHEAFLPFSIGKRVCLGEGAKAELELFFTTIIQAFLSLESPCPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAFLPFSIGKRVCLGEGAKAELELFFTTIIQAFLSLESPCPDPTLSLKPTVSG 480
QY 481 LFNIPPFAFQLOVRPTDLHSTTQTR 504
Db 481 LFNIPPFAFQLOVRPTDLHSTTQTR 504
RESULT 7
ABU66601
ID ABU66601 standard; protein; 504 AA.
XX
AC ABU66601;
XX
DT 23-MAY-2003 (first entry)
XX Human PRO polypeptide #32.
DE
XX
XX Human; PRO polypeptide; secreted and transmembrane protein;
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;
KW differentiation; chondrocyte; tumour; genetic disorder; cytostatic.
XX
OS Homo sapiens.
XX
XX US2003036180-A1.
XX
XX 20-FEB-2003.
XX
XX 09-MAY-2002; 2002US-00143114.
XX
XX 31-MAR-1997; 97WO-US0005230.
XX 12-JUN-1998; 98WO-US012456.
XX 14-JUL-1998; 98WO-US014552.
XX 28-AUG-1998; 98WO-US017888.
XX 10-SEP-1998; 98WO-US018824.
XX 14-SEP-1998; 98WO-US019093.
XX 14-SEP-1998; 98WO-US019094.
XX 14-SEP-1998; 98WO-US019177.
XX 16-SEP-1998; 98WO-US019330.
XX 17-SEP-1998; 98WO-US019437.
XX 07-OCT-1998; 98WO-US021141.
XX 29-OCT-1998; 98WO-US022991.
XX 29-OCT-1998; 98WO-US022992.
XX 20-NOV-1998; 98WO-US024855.
XX 01-DEC-1998; 98WO-US025108.
XX 05-JAN-1999; 99WO-US000106.
XX 08-MAR-1999; 99WO-US005028.
XX 10-MAR-1999; 99WO-US005190.
XX 20-APR-1999; 99WO-US008615.
XX 14-MAY-1999; 99WO-US010733.
XX 01-SEP-1999; 99WO-US012252.
XX 01-SEP-1999; 99WO-US020111.
XX 08-SEP-1999; 99WO-US020594.
XX 13-SEP-1999; 99WO-US020944.
XX 15-SEP-1999; 99WO-US021090.
XX 15-SEP-1999; 99WO-US021547.
XX 05-OCT-1999; 99WO-US023089.
XX 29-NOV-1999; 99WO-US028214.
XX 30-NOV-1999; 99WO-US028313.
XX 30-NOV-1999; 99WO-US028409.
XX 01-DEC-1999; 99WO-US028301.
XX 01-DEC-1999; 99WO-US028634.
XX 02-DEC-1999; 99WO-US028551.
XX 02-DEC-1999; 99WO-US028564.
XX 02-DEC-1999; 99WO-US028565.
XX 16-DEC-1999; 99WO-US030095.
XX 20-DEC-1999; 99WO-US030911.
XX 20-DEC-1999; 99WO-US030999.
XX 22-DEC-1999; 99WO-US030720.
XX 30-DEC-1999; 99WO-US031243.

PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000217.
 PR 06-JAN-2000; 2000WO-US000217.
 PR 11-FEB-2000; 2000WO-US000376.
 PR 18-FEB-2000; 2000WO-US0003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 22-FEB-2000; 2000WO-US004342.
 PR 24-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 01-MAR-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005746.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000US-00747259.
 PR 20-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001US-00796498.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 01-MAR-2001; 2001WO-US006666.
 PR 09-MAR-2001; 2001US-00802706.
 PR 14-MAR-2001; 2001US-00806889.
 PR 22-MAR-2001; 2001US-00816744.
 PR 05-APR-2001; 2001US-00828366.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854280.
 PR 18-MAY-2001; 2001US-00860216.
 PR 25-MAY-2001; 2001US-00866028.
 PR 25-MAY-2001; 2001US-00866034.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 21-JUN-2001; 2001US-00887879.
 PR 22-JUN-2001; 2001WO-US020116.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 18-JUL-2001; 2001US-00908827.
 PR 06-AUG-2001; 2001US-00924419.
 PR 09-AUG-2001; 2001US-00927796.
 PR 16-AUG-2001; 2001US-00931836.
 PR 19-DEC-2001; 2001US-00028072.
 (GETH) GENENTECH INC.

PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski RJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX WPI: 2003-332040/31.
 DR N-PSDB; ACA03634.
 XX
 PT New secreted and transmembrane PRO nucleic acids, useful for gene
 PT therapy, in chromosome and gene mapping, as chromosome markers, in tissue
 PT typing, and in chromosome identification.
 XX

PS Claim 12; Fig 64; 660pp; English.
 XX The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The PRO
 CC polypeptides are secreted and transmembrane proteins. The PRO
 CC polypeptides are useful for detecting other PRO polypeptides, for linking
 CC bioactive molecules to cells expressing PRO polypeptides, for modulating
 CC biological activities of cells expressing PRO polypeptides, and for for
 CC identifying agonists or antagonists. The PRO polypeptides are useful for
 CC for stimulating the release of tumour necrosis factor (TNF)-alpha from
 CC human blood, for stimulating the proliferation or differentiation of
 CC chondrocytes, and detecting the presence of tumours. The polynucleotide
 CC sequences encoding PRO polypeptides are useful as hybridisation probes,
 CC in chromosome and gene mapping, in the generation of antisense RNA and
 CC DNA, in the preparation of PRO polypeptides, for generating transgenic
 CC animals or knockout animals, for the genetic analysis of individuals with
 CC genetic disorders, and in gene therapy. ABU6570-ABU6684 represent the
 CC human PRO polypeptides of the invention. Note: The sequence data for this
 CC patent was obtained in electronic format directly from the USPTO web site
 CC at seqdata.uspto.gov/psipdbEntry.html
 XX
 SQ Sequence 504 AA;
 Query Match 100.0%; Score 2615; DB 6; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1.8e-258;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MEATGTWALLALALLLLTLLSGLTRARGHLPGPTPLPLGNNLQIRPGALYGLMRL 60
 Db 1 MEATGTWALLALALLLLTLLSGLTRARGHLPGPTPLPLGNNLQIRPGALYGLMRL 60
 Qy 61 SKYGPVFTIYLGPRPVVVLVQEAVERALGGQAEFSGRGTVAMLEGTFDGHGVFFSN 120
 Db 61 SKYGPVFTIYLGPRPVVVLVQEAVERALGGQAEFSGRGTVAMLEGTFDGHGVFFSN 120
 Qy 121 GERWQLRKFTMLALRDGLMGKREGEELIOAEARCLVETFGTEGRPFPSLLAQATSN 180
 Db 121 GERWQLRKFTMLALRDGLMGKREGEELIOAEARCLVETFGTEGRPFPSLLAQATSN 180
 Qy 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGHKQLLHH 240
 Db 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGHKQLLHH 240
 Qy 241 VSTLAFTVRQVQOQHGNLDASGPARDLVDAFLKWAQEEQNFTGNKMLMTVIYLL 300
 Db 241 VSTLAFTVRQVQOQHGNLDASGPARDLVDAFLKWAQEEQNFTGNKMLMTVIYLL 300
 Qy 301 FAGTMTVSTTVGYTLLLLMKYHVQKVRBELNRELGAGQAPSLGDRTRLPYTDAVLHEA 360
 Db 301 FAGTMTVSTTVGYTLLLLMKYHVQKVRBELNRELGAGQAPSLGDRTRLPYTDAVLHEA 360
 Qy 361 QRLLLVPMGIPRTLMTTRFRGYTLPGQTEVPFLGSLTHDPNIFKHPEENPDRFLDA 420
 Db 361 QRLLLVPMGIPRTLMTTRFRGYTLPGQTEVPFLGSLTHDPNIFKHPEENPDRFLDA 420
 Qy 421 DGRFRKHEAFLPFLSKGRVCLGEGAKAELFFFTTILQAFSLSPCPDPTSLKPTVSG 480
 Db 421 DGRFRKHEAFLPFLSKGRVCLGEGAKAELFFFTTILQAFSLSPCPDPTSLKPTVSG 480
 Qy 481 LFNIPPAFQQLVRPTDLHSTTQTR 504
 Db 481 LFNIPPAFQQLVRPTDLHSTTQTR 504
 RESULT 8
 ABUS9682
 ID ABUS9682 standard; protein; 504 AA.
 XX
 AC ABUS9682;
 XX
 DT 13-MAY-2003 (first entry)
 XX
 DE Novel secreted and transmembrane protein PRO1906.

XX Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumour; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
KW retinal neurons cell survival; rod photoreceptor cell survival;
KW retinal disorder; retinitis pigmentosa; kidney disease;
KW mammalian kidney mesangial cell proliferation; Berger disease;
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
KW chondrocyte redifferentiation; sports injury; arthritis.
XX
OS Homo sapiens.
XX
PN US2003017563-A1.
XX
XX 23-JAN-2003.
PD
XX 07-MAY-2002; 2002US-00140808.
XX
XX 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021144.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 98WO-US000106.
PR 08-MAR-1999; 98WO-US005028.
PR 10-MAR-1999; 98WO-US005190.
PR 20-APR-1999; 98WO-US008615.
PR 14-MAY-1999; 98WO-US010733.
PR 02-JUN-1999; 98WO-US013252.
PR 01-SEP-1999; 98WO-US020111.
PR 08-SEP-1999; 98WO-US020594.
PR 13-SEP-1999; 98WO-US020594.
PR 15-SEP-1999; 98WO-US021090.
PR 15-SEP-1999; 98WO-US021547.
PR 05-OCT-1999; 98WO-US023089.
PR 29-NOV-1999; 98WO-US028214.
PR 30-NOV-1999; 98WO-US028313.
PR 30-NOV-1999; 98WO-US028409.
PR 01-DEC-1999; 98WO-US028301.
PR 01-DEC-1999; 98WO-US028634.
PR 02-DEC-1999; 98WO-US028551.
PR 02-DEC-1999; 98WO-US028551.
PR 02-DEC-1999; 98WO-US028564.
PR 02-DEC-1999; 98WO-US028565.
PR 16-DEC-1999; 98WO-US030095.
PR 20-DEC-1999; 98WO-US030911.
PR 20-DEC-1999; 98WO-US030999.
PR 22-DEC-1999; 98WO-US030720.
PR 30-DEC-1999; 98WO-US031243.
PR 30-DEC-1999; 98WO-US031274.
PR 03-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004514.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2003-148238/14.
DR N-PSDB; ABX89172.
XX
XX Two hundred and seventy five nucleic acids encoding PRO polypeptides,
PT useful for treating pericyte-associated tumors, diabetes and various bone
PT and/or cartilage disorders, e.g. arthritis.
XX
XX Claim 12; Fig 64; 65pp; English.
XX
XX The invention describes an isolated human PRO polypeptide. The PRO
CC polypeptides are useful in detecting PRO polypeptides in a sample, in
CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and
CC in modulating at least one biological activity of a cell expressing a PRO
CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus
CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186
CC stimulate adrenal cortical capillary endothelial growth, and PRO536,
CC PRO943, PRO826, PRO1068 or PRO335, PRO826, PRO819, PRO1126,
CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus
CC useful for treating conditions or disorders where angiogenesis would be
CC beneficial, e.g. wound healing and antagonist of this polypeptide are

CC useful for treating cancerous tumours. PRO812 inhibits vascular
 CC endothelial growth factor (VEGF) stimulated proliferation of endothelial
 CC cells and is thus useful for inhibiting endothelial cell growth in
 CC mammals which would be beneficial in inhibiting tumour growth. PRO826,
 CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of
 CC stimulated T-lymphocytes and are therapeutically useful for enhancing
 CC immune response. PRO828, PRO826, PRO1068 or PRO132 enhance survival of
 CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of
 CC rod photoreceptor cells) and therefore are useful for treating retinal
 CC disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813
 CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,
 CC and therefore are useful for treating kidney disorders associated with
 CC decreased mesangial cell function such as Berger disease or other
 CC nephropathies associated with dermatitis, herpetiformis or Crohn's
 CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the
 CC proliferation and/or redifferentiation of chondrocytes in culture and are
 CC thus useful for treating sports injuries, and arthritis. This is the
 CC amino acid sequence of a novel human PRO protein
 XX Sequence 504 AA;

Query Match 100.0%; Score 2615; DB 6; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1.8e-258; Indels 0; Gaps 0;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLLTALSGTRARGLPPGPTPLPLGLNLLQRLPGALYGLMRL 60
 DB 1 MEATGTWALLALALLLLTALSGTRARGLPPGPTPLPLGLNLLQRLPGALYGLMRL 60

QY 61 SKKYGVPFTIYLGWPRVWVLVGEAVREALGGQAEFSGRTVAMLEGTFDGHGVFFSN 120
 DB 61 SKKYGVPFTIYLGWPRVWVLVGEAVREALGGQAEFSGRTVAMLEGTFDGHGVFFSN 120

QY 121 GERWRQLRKFTMLALRDLGMKRGEEELIQAEARCLVETFGTEGRPDPSLLLAQATSN 180
 DB 121 GERWRQLRKFTMLALRDLGMKRGEEELIQAEARCLVETFGTEGRPDPSLLLAQATSN 180

QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
 DB 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAFTVROVQHQGNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300
 DB 241 VSTLAFTVROVQHQGNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYLL 300

QY 301 FACTMTVSTTVGYTLMLLMKYPHVQKWVRELARELGAGQAPSLGDRTRLPYTDVULHEA 360
 DB 301 FACTMTVSTTVGYTLMLLMKYPHVQKWVRELARELGAGQAPSLGDRTRLPYTDVULHEA 360

QY 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIPKHPEEFPNDRFLDA 420
 DB 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIPKHPEEFPNDRFLDA 420

QY 421 DGRFRKHEAFPLPSLGRVCLGEGAKAEFLPFTTILQAFSLESPCPDPTLSLKPTVSG 480
 DB 421 DGRFRKHEAFPLPSLGRVCLGEGAKAEFLPFTTILQAFSLESPCPDPTLSLKPTVSG 480

QY 481 LFNIPPAFQLQVRPTDLHSTTTQTR 504
 DB 481 LFNIPPAFQLQVRPTDLHSTTTQTR 504

RESULT 9

ABO24872
 ID ABO24872 standard; protein; 504 AA.
 XX
 AC ABO24872;
 XX
 DT 05-SEP-2003 (first entry)
 DE Human secreted/transmembrane protein (PRO) #32.
 XX
 KW Human; PRO; secreted protein; transmembrane protein; tumour; cytostatic;

KW gene therapy; tumour necrosis factor-alpha; TNF-alpha; blood;
 KW proteoglycan; cartilage; cytokine; peripheral blood mononuclear cell;
 KW PBMC; Glucose uptake; FFA; skeletal muscle cell; adipocyte cell;
 KW chondrocyte cell proliferation; chondrocyte cell differentiation;
 KW pericyte cell; inner ear utricular supporting cell; T-lymphocyte cell;
 KW endothelial cell; A-peptide; factor VIIA.
 XX Homo sapiens.
 OS
 PN US2003036179-A1.
 PD 20-FEB-2003.
 XX
 PF 10-MAY-2002; 2002US-00142431.
 XX
 PR 31-MAR-1997; 97WO-US005230.
 PR 12-JUN-1998; 98WO-US012456.
 PR 14-JUL-1998; 98WO-US014552.
 PR 28-AUG-1998; 98WO-US017888.
 PR 10-SEP-1998; 98WO-US018824.
 PR 14-SEP-1998; 98WO-US019093.
 PR 14-SEP-1998; 98WO-US019094.
 PR 14-SEP-1998; 98WO-US019177.
 PR 16-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98WO-US019437.
 PR 07-OCT-1998; 98WO-US021141.
 PR 29-OCT-1998; 98WO-US022991.
 PR 29-OCT-1998; 98WO-US022992.
 PR 20-NOV-1998; 98WO-US024855.
 PR 01-DEC-1998; 98WO-US025108.
 PR 05-JAN-1999; 99WO-US000106.
 PR 08-MAR-1999; 99WO-US005028.
 PR 10-MAR-1999; 99WO-US005190.
 PR 20-APR-1999; 99WO-US008615.
 PR 14-MAY-1999; 99WO-US010733.
 PR 02-JUN-1999; 99WO-US012252.
 PR 01-SEP-1999; 99WO-US020111.
 PR 08-SEP-1999; 99WO-US020594.
 PR 13-SEP-1999; 99WO-US020944.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.
 PR 03-OCT-1999; 99WO-US023089.
 PR 23-NOV-1999; 99WO-US028214.
 PR 30-NOV-1999; 99WO-US028313.
 PR 30-NOV-1999; 99WO-US028409.
 PR 01-DEC-1999; 99WO-US028301.
 PR 01-DEC-1999; 99WO-US028634.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028564.
 PR 16-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 22-DEC-1999; 99WO-US030999.
 PR 30-DEC-1999; 99WO-US030720.
 PR 30-DEC-1999; 99WO-US031243.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 01-MAR-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005746.
 PR 10-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006319.
 PR 20-MAR-2000; 2000WO-US006894.
 PR 21-MAR-2000; 2000WO-US007377.
 PR 30-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.

17-MAY-2000; 2000WO-US013705.
 22-MAY-2000; 2000WO-US014042.
 30-MAY-2000; 2000WO-US014941.
 02-JUN-2000; 2000WO-US015264.
 28-JUL-2000; 2000WO-US020710.
 11-AUG-2000; 2000WO-US022031.
 23-AUG-2000; 2000WO-US023522.
 24-AUG-2000; 2000WO-US023328.
 08-NOV-2000; 2000WO-US030952.
 10-NOV-2000; 2000WO-US030873.
 01-DEC-2000; 2000WO-US032678.
 20-DEC-2000; 2000US-00747259.
 20-DEC-2000; 2000WO-US034956.
 28-FEB-2001; 2001US-00796498.
 28-FEB-2001; 2001WO-US006520.
 01-MAR-2001; 2001WO-US006666.
 09-MAR-2001; 2001US-00802706.
 14-MAR-2001; 2001US-00808689.
 22-MAR-2001; 2001US-00816744.
 05-APR-2001; 2001US-00828366.
 10-MAY-2001; 2001US-00854208.
 10-MAY-2001; 2001US-00854280.
 18-MAY-2001; 2001US-00860216.
 25-MAY-2001; 2001US-00866028.
 25-MAY-2001; 2001US-00866034.
 25-MAY-2001; 2001WO-US017092.
 01-JUN-2001; 2001US-00872035.
 01-JUN-2001; 2001WO-US017800.
 05-JUN-2001; 2001US-00874503.
 14-JUN-2001; 2001US-00882636.
 19-JUN-2001; 2001US-00886342.
 20-JUN-2001; 2001WO-US019692.
 21-JUN-2001; 2001US-00887879.
 22-JUN-2001; 2001WO-US020116.
 29-JUN-2001; 2001WO-US021066.
 09-JUL-2001; 2001US-00921735.
 18-JUL-2001; 2001US-00908627.
 06-AUG-2001; 2001US-00924419.
 09-AUG-2001; 2001US-00927796.
 16-AUG-2001; 2001US-00931836.
 19-DEC-2001; 2001US-00028072.

(GETH) GENENTECH INC.

Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 WPT; 2003-466355/44.
 N-PSDB; ACD41826.

New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO1114 or
 PRO4978, useful in molecular biology, chromosome and gene mapping, in
 generating antisense RNA and DNA, and in gene therapy.

Claim 12; Fig 64; 659pp; English.

The invention relates to an isolated nucleic acid comprising at least 80%
 sequence identity to a PRO (secreted and transmembrane protein) cDNA
 sequence identity to a PRO polypeptide, or its
 comprising a nucleic acid (a) encoding a PRO polypeptide, or its
 extracellular domain (with or without its associated signal peptide),
 which comprises any of the 275 120-850 residue amino acid sequences,
 given in the specification; (b) comprising any of the 275 300-3500
 nucleotide sequences, given in the specification; or (c) comprising the
 full-length coding sequence of the nucleotide sequences given in the
 specification, or of the DNA deposited under any of the American Type
 Culture Collection (ATCC) Accession Numbers listed in the specification.
 Also included are a vector comprising the novel nucleic acid, a host cell
 comprising the vector, producing a PRO polypeptide, the isolated PRO
 polypeptides detailed above, a chimeric molecule comprising the PRO
 polypeptide of fused to a heterologous amino acid sequence, an anti-PRO
 antibody, detecting a PRO polypeptide in a sample suspected of containing
 the PRO polypeptide, linking a bioactive molecule to a cell expressing a

DT 27-MAY-2003 (first entry)
 DE Human secreted/transmembrane, PRO, protein SEQ ID 64.
 XX
 KW Human; secreted protein; transmembrane protein; PRO;
 KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;
 KW infertility; birth defects; premature aging; AIDS; biosensor;
 KW acquired immunodeficiency syndrome; cancer; diabetic complication;
 KW bioreactor; tumour.
 XX
 OS Homo sapiens.
 XX
 XX US2003032155-A1.
 PN
 PD 13-FEB-2003.
 XX
 XX 03-MAY-2002; 2002US-00137865.
 XX
 PR 31-MAR-1997; 97WO-US005230.
 PR 12-JUN-1998; 98WO-US012456.
 PR 14-JUL-1998; 98WO-US014552.
 PR 28-AUG-1998; 98WO-US017888.
 PR 10-SEP-1998; 98WO-US018824.
 PR 14-SEP-1998; 98WO-US019033.
 PR 14-SEP-1998; 98WO-US019094.
 PR 14-SEP-1998; 98WO-US019177.
 PR 16-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98WO-US019437.
 PR 07-OCT-1998; 98WO-US021141.
 PR 29-OCT-1998; 98WO-US022991.
 PR 29-OCT-1998; 98WO-US022992.
 PR 20-NOV-1998; 98WO-US024855.
 PR 01-DEC-1998; 98WO-US025108.
 PR 05-JAN-1999; 99WO-US000106.
 PR 08-MAR-1999; 99WO-US005028.
 PR 10-MAR-1999; 99WO-US005190.
 PR 20-APR-1999; 99WO-US008615.
 PR 14-MAY-1999; 99WO-US010733.
 PR 02-JUN-1999; 99WO-US012252.
 PR 01-SEP-1999; 99WO-US020111.
 PR 08-SEP-1999; 99WO-US020594.
 PR 13-SEP-1999; 99WO-US020944.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.
 PR 03-OCT-1999; 99WO-US023089.
 PR 29-NOV-1999; 99WO-US028214.
 PR 30-NOV-1999; 99WO-US028313.
 PR 30-NOV-1999; 99WO-US028409.
 PR 01-DEC-1999; 99WO-US028301.
 PR 01-DEC-1999; 99WO-US028634.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028564.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 20-DEC-1999; 99WO-US030999.
 PR 22-DEC-1999; 99WO-US030720.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005041.
 PR 02-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005746.
 PR 10-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006319.
 PR 20-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000US-00747259.
 PR 20-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001US-00796498.
 PR 28-FEB-2001; 2001US-00796498.
 PR 01-MAR-2001; 2001WO-US006520.
 PR 09-MAR-2001; 2001US-00802706.
 PR 14-MAR-2001; 2001US-00808689.
 PR 22-MAR-2001; 2001US-00816744.
 PR 05-APR-2001; 2001US-00828366.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854208.
 PR 18-MAY-2001; 2001US-00860216.
 PR 25-MAY-2001; 2001US-00866028.
 PR 25-MAY-2001; 2001US-00866034.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 21-JUN-2001; 2001US-00887879.
 PR 22-JUN-2001; 2001WO-US020116.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 18-JUL-2001; 2001US-00908827.
 PR 06-AUG-2001; 2001US-00924419.
 PR 09-AUG-2001; 2001US-00927796.
 PR 16-AUG-2001; 2001US-00931836.
 PR 19-DEC-2001; 2001US-00028072.
 XX
 PA (GETH) GENENTECH INC.
 XX
 XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI; 2003-331925/31.
 DR N-PSDB; ACA04055.
 XX
 PT New secreted and transmembrane nucleic acids and polypeptides, designated
 as PRO, useful for treating inflammation, organ failure, atherosclerosis,
 cardiac injury, infertility, birth defects, premature aging, AIDS, or
 cancer.
 PT
 PT
 XX Claim 12; Fig 64; 659pp; English.
 PS
 XX The invention relates to an isolated nucleic acid comprising, or which is
 at least 80% identical to, or the full-length coding sequence of, any of
 the 275 nucleotide sequences, encoding the corresponding PRO polypeptide
 (one of 275 secreted or transmembrane proteins). The nucleic acid further
 comprises the full-length coding sequence of the DNA deposited under
 American Type Culture Collection (ATCC) accession number in a list given
 in the specification. Also included are vectors and host cells for
 producing PRO proteins. PRO fusion proteins, anti-PRO antibodies, PRO
 extracellular domains and mature sequences, methods of detecting PRO
 proteins, methods for stimulating the release of TNF-alpha (tumour
 necrosis factor alpha) from human blood, (and the proliferation of
 differentiation of chondrocyte cells, the proliferation of, or gene

Mon Sep 20 09:09:24 2004

CC expression in pericyte cells, the release or proteoglycans from
 CC cartilage, proliferation of inner ear utricular supporting cells, the
 CC proliferation of T-lymphocyte cells, the release of a cytokine from
 CC peripheral blood mononuclear cells (PBMC), or the proliferation of
 CC endothelial cells), a method for modulating the uptake of glucose or free
 CC fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the
 CC binding of A-peptide to factor VIIA, or the differentiation of adipocyte
 CC cells, a method for detecting the presence of a tumour in a mammal and an
 CC oligonucleotide probe derived from any of the nucleotide sequences cited
 CC above. The nucleic acids and polypeptides are useful for treating
 CC inflammatory diseases, organ failure, atherosclerosis, cardiac injury,
 CC infertility, birth defects, premature aging, AIDS (acquired
 CC immunodeficiency syndrome), cancer, or diabetic complications. The
 CC nucleic acids are useful as hybridisation probes, in chromosome and gene
 CC mapping, and in generating antisense RNA or DNA. The polypeptides are
 CC useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both
 CC are useful in tissue typing. The present sequence represents a PRO
 CC protein of the invention
 XX
 SQ Sequence 504 AA;

Query Match 100.0%; Score 2615; DB 6; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1.8e-258;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTWALLALALLLTALSGTRARGLPPGPTPLPLGNLLQLRPGALYGLMRL 60
 Db 1 MEATGTWALLALALLLTALSGTRARGLPPGPTPLPLGNLLQLRPGALYGLMRL 60
 QY 61 SKKYGPFVYIYIGMPWPVVVLVGQEAVERALGQAEFSGRGTVAMLEGTDFGQGVFFN 120
 Db 61 SKKYGPFVYIYIGMPWPVVVLVGQEAVERALGQAEFSGRGTVAMLEGTDFGQGVFFN 120
 QY 121 GERWRQLRFTMLALDLGKREGBELQAEARCLIVETFOGTGPRFPDPSLLAQATSN 180
 Db 121 GERWRQLRFTMLALDLGKREGBELQAEARCLIVETFOGTGPRFPDPSLLAQATSN 180
 QY 181 VVCSLLFGLRFSYEDKEFQAVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
 Db 181 VVCSLLFGLRFSYEDKEFQAVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
 QY 241 VSTLAFTVQVQHQHNLDAAGPARLDVDAFLLKMAEQEONPTEFTNKNMLMTVYLL 300
 Db 241 VSTLAFTVQVQHQHNLDAAGPARLDVDAFLLKMAEQEONPTEFTNKNMLMTVYLL 300
 QY 301 FAGTMTVSTTVGYTLALLMKYPHVQKWRBELNRELGAQAPSLGDRTRLPYTDVILHEA 360
 Db 301 FAGTMTVSTTVGYTLALLMKYPHVQKWRBELNRELGAQAPSLGDRTRLPYTDVILHEA 360
 QY 361 QRLIALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIFKHPEEFNPDRLDA 420
 Db 361 QRLIALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIFKHPEEFNPDRLDA 420
 QY 421 DGRPRKEAFPLPSIGKRVCLGEGAKAELELFTTILQAFSLSPCPDPTLSLKPTVSG 480
 Db 421 DGRPRKEAFPLPSIGKRVCLGEGAKAELELFTTILQAFSLSPCPDPTLSLKPTVSG 480
 QY 481 LFNIPPAFQLOVRPDLHSTTQTR 504
 Db 481 LFNIPPAFQLOVRPDLHSTTQTR 504

RESULT 11
 ADA45583
 ID ADA45583 standard; protein; 504 AA.
 XX
 AC ADA45583;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO1906.
 XX
 KW Human; secreted and transmembrane protein; PRO;

KW Tumour necrosis factor alpha release; TNF-alpha release;
 KW glucose uptake modulator; FFA uptake modulator;
 KW cell proliferation stimulator; cell differentiation stimulator;
 KW cell differentiation inhibitor; cytokine release stimulator; tumour;
 KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW cervical tumour; liver tumour; chromosome mapping; gene mapping;
 KW gene therapy; chromosome identification; chromosome marker.
 XX Homo sapiens.
 OS
 PN US2003022328-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 16-APR-2002; 2002US-00123904.
 PR 31-MAR-1997; 97WO-US005230.
 PR 12-JUN-1998; 98WO-US012456.
 PR 14-JUL-1998; 98WO-US014552.
 PR 28-AUG-1998; 98WO-US017888.
 PR 10-SEP-1998; 98WO-US018824.
 PR 14-SEP-1998; 98WO-US019093.
 PR 14-SEP-1998; 98WO-US019094.
 PR 14-SEP-1998; 98WO-US019177.
 PR 16-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98WO-US019437.
 PR 07-OCT-1998; 98WO-US021141.
 PR 29-OCT-1998; 98WO-US022991.
 PR 29-OCT-1998; 98WO-US022992.
 PR 20-NOV-1998; 98WO-US024855.
 PR 01-DEC-1998; 98WO-US025108.
 PR 05-JAN-1999; 99WO-US000106.
 PR 08-MAR-1999; 99WO-US005028.
 PR 10-MAR-1999; 99WO-US005190.
 PR 20-APR-1999; 99WO-US008615.
 PR 14-MAY-1999; 99WO-US010733.
 PR 02-JUN-1999; 99WO-US01252.
 PR 01-SEP-1999; 99WO-US020111.
 PR 08-SEP-1999; 99WO-US020594.
 PR 13-SEP-1999; 99WO-US020944.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.
 PR 05-OCT-1999; 99WO-US023089.
 PR 29-NOV-1999; 99WO-US028214.
 PR 30-NOV-1999; 99WO-US028313.
 PR 30-NOV-1999; 99WO-US028409.
 PR 01-DEC-1999; 99WO-US028301.
 PR 01-DEC-1999; 99WO-US028634.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028564.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 20-DEC-1999; 99WO-US030999.
 PR 30-DEC-1999; 99WO-US030720.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005746.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 21-MAR-2000; 2000WO-US007532.

PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 01-DEC-2000; 2000WO-US030873.
 PR 10-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000US-00747259.
 PR 28-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001US-00796498.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 01-MAR-2001; 2001WO-US006666.
 PR 09-MAR-2001; 2001US-00802706.
 PR 14-MAR-2001; 2001US-00808699.
 PR 22-MAR-2001; 2001US-00816744.
 PR 05-APR-2001; 2001US-00828366.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854280.
 PR 18-MAY-2001; 2001US-00860216.
 PR 25-MAY-2001; 2001US-00866028.
 PR 25-MAY-2001; 2001US-00866034.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 21-JUN-2001; 2001US-00887879.
 PR 22-JUN-2001; 2001WO-US020116.
 PR 29-JUN-2001; 2001WO-US021056.
 PR 03-JUL-2001; 2001WO-US021735.
 PR 18-JUL-2001; 2001US-00908827.
 PR 06-AUG-2001; 2001US-00924419.
 PR 09-AUG-2001; 2001US-00927796.
 PR 16-AUG-2001; 2001US-00931836.
 PR 19-DEC-2001; 2001US-00028072.
 (GETH) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX WPI; 2003-584997/55.
 DR N-PSDB; ADA45582.
 DR
 XX Novel secreted and transmembrane polypeptide for modulating biological
 PT activity of cell expressing the polypeptide, identifying agonists or
 PT antagonists of polypeptide, and as molecular weight markers.
 XX
 XX Claim 12; Fig 64; 659pp; English.

CC The invention describes 305 nucleic acids encoding PRO (secreted and
 CC transmembrane) polypeptides (I). (I) is useful for stimulating the
 CC release of TNF-alpha from human blood, for modulating the uptake of
 CC glucose or FFA by skeletal muscle cells or adipocyte cells, for
 CC stimulating the proliferation or differentiation of chondrocyte cells,
 CC cells, for stimulating the proliferation of or gene expression in pericyte
 CC cells, for stimulating the release of proteoglycans from cartilage, for
 CC stimulating the proliferation of inner ear utricular supporting cells,
 CC for stimulating the proliferation of T-lymphocyte cells, for stimulating
 CC the release of a cytokine from PBMC cells, for inhibiting the binding of
 CC A-peptide to factor VIIA, for inhibiting the differentiation of adipocyte
 CC cells, for stimulating proliferation of endothelial cells, for detecting
 CC the presence of tumour in a mammal. The tumour is lung, colon, breast,
 CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes
 CC are useful for isolating genomic and cDNA nucleotide sequences or

CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful
 CC in assays to identify other proteins or molecules involved in binding
 CC interaction. A polynucleotide (II) encoding (I) is useful in chromosome
 CC and gene mapping, in generation of antisense RNA and DNA, in the
 CC preparation of PRO polypeptide, for generating transgenic animals or
 CC knockout animals which in turn are useful in the development and
 CC screening of therapeutically useful reagents, in gene therapy, for
 CC chromosome identification, as chromosome marker, and for generating
 CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.
 CC detecting its expression in specific cells, tissues or serum, and for
 CC affinity purification of PRO from recombinant cell culture or natural
 CC sources. (I) and (II) are useful for tissue typing. This is the amino
 CC acid sequence of a novel human secreted and transmembrane PRO
 CC polypeptide.

XX SQ Sequence 504 AA;

Query Match 100.0%; Score 2615; DB 6; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1.8e-258;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MEATGTWALLLALALLLLLTALSGTRARGHLPGPTPLGLNLLQLRPGALYSGLMRL	60
Db	1	MEATGTWALLLALALLLLLTALSGTRARGHLPGPTPLGLNLLQLRPGALYSGLMRL	60
QY	61	SKYGPVFTIYLGPMRPVVLVQEAQVREALGGQAEFFSGRTVAMLEGTDFDGHGVFFSN	120
Db	61	SKYGPVFTIYLGPMRPVVLVQEAQVREALGGQAEFFSGRTVAMLEGTDFDGHGVFFSN	120
QY	121	GERWQLRKETMLALRDLGMKGEGEBELIQAEARCLVETFGTEGRFPDPSLLAQATSN	180
Db	121	GERWQLRKETMLALRDLGMKGEGEBELIQAEARCLVETFGTEGRFPDPSLLAQATSN	180
QY	181	VVCSLLFGLRFSYEDKEFOAVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH	240
Db	181	VVCSLLFGLRFSYEDKEFOAVRAAGGTLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH	240
QY	241	VSTLAFTVRQVQHQGNLDASGPARDLVDAFLLLKMAQESQNPCTFTNKNMLMTVIYLL	300
Db	241	VSTLAFTVRQVQHQGNLDASGPARDLVDAFLLLKMAQESQNPCTFTNKNMLMTVIYLL	300
QY	301	FAGTMVSTIVGYTILLMLKYPHVQKVRRELNRELGAGQAPSLGDRTRLPYTDVILHEA	360
Db	301	FAGTMVSTIVGYTILLMLKYPHVQKVRRELNRELGAGQAPSLGDRTRLPYTDVILHEA	360
QY	361	QRLIALVPMGIPRTLMRTTFRGTYLPGQTEVPLGSLIHDNIFKHPEEPNDRFLDA	420
Db	361	QRLIALVPMGIPRTLMRTTFRGTYLPGQTEVPLGSLIHDNIFKHPEEPNDRFLDA	420
QY	421	DGRFRKHEAFPLPSLGKRVCLGEGLAKELEFVFTTILQAFSLESPPDPTLSLKPTVSG	480
Db	421	DGRFRKHEAFPLPSLGKRVCLGEGLAKELEFVFTTILQAFSLESPPDPTLSLKPTVSG	480
QY	481	LFNIPPAFOLQVRPTDLHSTTQTR	504
Db	481	LFNIPPAFOLQVRPTDLHSTTQTR	504

RESULT 12

ADA76014
 ID ADA76014 standard; protein; 504 AA.

XX ADA76014;

XX 20-NOV-2003 (first entry)

DE Human PRO polypeptide #32.

XX Human; PRO; secreted polypeptide; transmembrane polypeptide;
 KW tumour necrosis factor-alpha; TNF-alpha; chondrocyte cell; tumour;
 KW cancer; adrenal; lung; colon; breast; prostate; kidney; cervix;
 KW liver; microvascular endothelial cell; glucose; FFA;
 KW skeletal muscle cell; adipocyte cell; pericyte cell;

KW inner ear utricular supporting cell; T-lymphocyte cell;
KW endothelial cell tube formation; bone disorder; cartilage disorder;
KW sports injury; proteoglycan; articular cartilage defect; osteoarthritis;
KW rheumatoid arthritis; haemoglobin-associated disorder thalassaemia;
KW immune system cell infiltration.
XX
OS Homo sapiens.
XX
XX US2003073212-A1.
XX
XX PD 17-APR-2003.
XX
XX PF 16-APR-2002; 2002US-00123903.
XX
XX 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000108.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUN-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-0074259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
XX (GETH) GENENTECH INC.
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2003-687639/65.
DR N-PSDB; ADA76013.
XX
XX New isolated nucleic acid encoding a secreted and transmembrane
PT polypeptide, designated e.g. PRO1114 or PRO4978, useful in chromosome and
PT gene mapping, in generating antisense RNA and DNA, and in gene therapy.
XX
PS Claim 12; Fig 64; 659pp; English.
XX
XX The invention relates to isolated human PRO polypeptides (secreted and
CC transmembrane polypeptides) and the polynucleotides encoding them. The
CC invention also relates to an antibody which specifically binds to a PRO
CC polypeptide, a method for stimulating the release of tumour necrosis
CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells and a method for
CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung, the
CC colon, breast, prostate, rectal, kidney, cervical and liver tumours). The
CC polynucleotides are useful in molecular biology, including uses as
CC hybridisation probes, in chromosome and gene mapping, in generating
CC antisense RNA and DNA and in gene therapy. The polynucleotides may also
CC be used in preparing PRO polypeptides by recombinant techniques and in
CC generating either transgenic animals or knock-out animals which are
CC useful in the development and screening of therapeutically useful
CC reagents. The PRO polypeptides or antibodies are used in preparing a
CC medicament for treating a condition responsive to the polypeptides or
CC antibodies, such as tumours, for stimulating and inhibiting proliferation

CC of human microvascular endothelial cells, for modulating the uptake of
CC glucose or FFA by skeletal muscle cells or adipocyte cells, for
CC stimulating differentiation of adipocyte cells, for stimulating
CC proliferation of or gene expression in pericyte cells, for stimulating
CC the proliferation of inner ear utricular supporting cells or T-lymphocyte
CC cells, for inducing endothelial cell tube formation and for treating
CC various bone and/or cartilage disorders such as sports injuries and
CC arthritis. PRO polypeptides which stimulate the release of proteoglycans
CC from cartilage are useful for treating sports-related joint problems, PRO
CC articular cartilage defects, osteoarthritis and rheumatoid arthritis. PRO
CC polypeptides are also useful for treating various mammalian haemoglobin-
CC associated disorders such as various thalassaemias and conditions which
CC may benefit from enhanced local immune system cell infiltration. This
CC sequence represents a human PRO polypeptide of the invention. Note: The
CC sequence data for this patent is also available in electronic format from
CC USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ

Sequence 504 AA;

Query Match 100.0%; Score 2615; DB 6; Length 504;

Best Local Similarity 100.0%; Pred. No. 1, de-258;

Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATGTTWALLLALLLTLALSGTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRL 60
DB 1 MEATGTTWALLLALLLTLALSGTRARGHLPGPTPLPLGNNLQLRPGALYSGLMRL 60
QY 61 SKKYGPFVFTYLGWPWRPVVLVGGAEVREALGGQAEFGSGRTVAMLEGTFDGHGVFFSN 120
DB 61 SKKYGPFVFTYLGWPWRPVVLVGGAEVREALGGQAEFGSGRTVAMLEGTFDGHGVFFSN 120
QY 121 GERWRQLRKFTMLALRDLMGKREGEELIQAEARCLVETFGTEGRFPDPSLLLAQATSN 180
DB 121 GERWRQLRKFTMLALRDLMGKREGEELIQAEARCLVETFGTEGRFPDPSLLLAQATSN 180
QY 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRFSYEDKEFQAVVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAFTVQVQHQNLDSAGPARDLVDAFLIKMAQEQNPGTEFTNKNMLMTVIYLL 300
DB 241 VSTLAFTVQVQHQNLDSAGPARDLVDAFLIKMAQEQNPGTEFTNKNMLMTVIYLL 300
QY 301 FAGTMVSTTVGTYLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTPYTDVILHEA 360
DB 301 FAGTMVSTTVGTYLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTPYTDVILHEA 360
QY 361 ORLLALVPMGIPRTLMRTTRFRGYTLPQGTVEVPLLGSLILHDPNIFKHPEEFNDRFLDA 420
DB 361 ORLLALVPMGIPRTLMRTTRFRGYTLPQGTVEVPLLGSLILHDPNIFKHPEEFNDRFLDA 420
QY 421 DGRFRKHEAPLPSLCKRVCGLGSLAKAELELFFFTTILQAFSLSECPDPTLSLKTQVSG 480
DB 421 DGRFRKHEAPLPSLCKRVCGLGSLAKAELELFFFTTILQAFSLSECPDPTLSLKTQVSG 480
QY 481 LFNIPPAFQLQVRPTDLHSTTQTR 504
DB 481 LFNIPPAFQLQVRPTDLHSTTQTR 504

RESULT 13

ADA18664

ID ADA18664 standard; protein; 504 AA.

XX

AC ADA18664;

XX

DT 20-NOV-2003 (first entry)

XX

DE Human PRO polypeptide #32.

XX

KW Human; PRO; secreted polypeptide; transmembrane polypeptide;
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; lung;
KW colon; breast; prostate; rectum; cervix; liver; tumour; cancer;KW glucose uptake; FFA; adipocyte cell; pericyte cell; proteoglycan;
KW cartilage; inner ear utricular supporting cell; cytokine; A-peptide;
KW factor VIIA; endothelial cell.
OS Homo sapiens.

XX US2003054517-A1.

PN 20-MAR-2003.

PD 08-MAY-2002; 2002US-00141755.

PF 31-MAR-1997; 97WO-US005230.

XX 12-JUN-1998; 98WO-US012456.

PR 14-JUL-1998; 98WO-US014552.

PR 28-AUG-1998; 98WO-US017888.

PR 10-SEP-1998; 98WO-US018824.

PR 14-SEP-1998; 98WO-US019093.

PR 14-SEP-1998; 98WO-US019094.

PR 16-SEP-1998; 98WO-US019177.

PR 17-SEP-1998; 98WO-US019437.

PR 07-OCT-1998; 98WO-US021141.

PR 29-OCT-1998; 98WO-US022991.

PR 29-OCT-1998; 98WO-US022992.

PR 20-NOV-1998; 98WO-US024855.

PR 01-DEC-1998; 98WO-US025108.

PR 05-JAN-1999; 98WO-US00106.

PR 08-MAR-1999; 98WO-US005028.

PR 10-MAR-1999; 98WO-US005190.

PR 20-APR-1999; 98WO-US008615.

PR 14-MAY-1999; 98WO-US010733.

PR 02-JUN-1999; 98WO-US012252.

PR 01-SEP-1999; 98WO-US020111.

PR 08-SEP-1999; 98WO-US020594.

PR 13-SEP-1999; 98WO-US020944.

PR 15-SEP-1999; 98WO-US021090.

PR 15-SEP-1999; 98WO-US021547.

PR 05-OCT-1999; 98WO-US023089.

PR 29-NOV-1999; 98WO-US028214.

PR 30-NOV-1999; 98WO-US028313.

PR 30-NOV-1999; 98WO-US028409.

PR 01-DEC-1999; 98WO-US028301.

PR 01-DEC-1999; 98WO-US028634.

PR 02-DEC-1999; 98WO-US028551.

PR 02-DEC-1999; 98WO-US028564.

PR 02-DEC-1999; 98WO-US028565.

PR 16-DEC-1999; 98WO-US030095.

PR 20-DEC-1999; 98WO-US030911.

PR 22-DEC-1999; 98WO-US030999.

PR 30-DEC-1999; 98WO-US030720.

PR 30-DEC-1999; 98WO-US031243.

PR 30-DEC-1999; 98WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000277.

PR 06-JAN-2000; 2000WO-US000376.

PR 11-FEB-2000; 2000WO-US003565.

PR 18-FEB-2000; 2000WO-US004341.

PR 18-FEB-2000; 2000WO-US004342.

PR 22-FEB-2000; 2000WO-US004414.

PR 24-FEB-2000; 2000WO-US004914.

PR 24-FEB-2000; 2000WO-US005004.

PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005746.

PR 02-MAR-2000; 2000WO-US005841.

PR 10-MAR-2000; 2000WO-US006319.

PR 15-MAR-2000; 2000WO-US006884.

PR 20-MAR-2000; 2000WO-US007377.

PR 21-MAR-2000; 2000WO-US007532.

PR 30-MAR-2000; 2000WO-US008439.

PR 17-MAY-2000; 2000WO-US013705.

PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.

CC supporting cells, for stimulating the release of cytokines from PBMC
CC cells, for inhibiting the binding of A-peptide to factor VIIA, for
CC inhibiting the differentiation of adipocyte cells and for stimulating the
CC proliferation of endothelial cells. This sequence represents a human PRO
CC polypeptide of the invention. Note: The sequence data for this patent is
CC also available in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 504 AA;

Query Match 100.0%; Score 2615; DB 6; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.8e-258; Indels 0; Gaps 0;
Matches 504; Conservative 0; Mismatches 0;

QY 1 MEATGTWALLALALLLLTLLALSGTRARGHLPGPTPLLLGNLQLRPGALYGLMRL 60
DB 1 MEATGTWALLALALLLLTLLALSGTRARGHLPGPTPLLLGNLQLRPGALYGLMRL 60

QY 61 SKKYGPFVTIYLGPRPVVVLVGQAVREALGGQAEFSGRTVAMLEGTDFGHGVFFSN 120
DB 61 SKKYGPFVTIYLGPRPVVVLVGQAVREALGGQAEFSGRTVAMLEGTDFGHGVFFSN 120

QY 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOCTEGRPPDPESLLLAQATSN 180
DB 121 GERWQLRKFTMLALRDLMGKREGEELIQAEARCLVETFOCTEGRPPDPESLLLAQATSN 180

QY 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSGGQTYEMPSWFLRPLPGPHKQLLHH 240
DB 181 VVCSLLFGLRFSYEDKEFOAVVRAAGTLLGVSSGGQTYEMPSWFLRPLPGPHKQLLHH 240

QY 241 VSTLAAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEEQNGTEFTNKNNMLMTVIYLL 300
DB 241 VSTLAAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEEQNGTEFTNKNNMLMTVIYLL 300

QY 301 FAGTMTVSTVGYTLLMLMKYDHVKVREELNRELGAGQAPSLGDRTRLPYTDVAVLHEA 360
DB 301 FAGTMTVSTVGYTLLMLMKYDHVKVREELNRELGAGQAPSLGDRTRLPYTDVAVLHEA 360

QY 361 QRLALLVPMGIPRTLMRTTRFRGYTLPGQTEVFPPLLSILHDPNIFKGFEEFNPDRLDA 420
DB 361 QRLALLVPMGIPRTLMRTTRFRGYTLPGQTEVFPPLLSILHDPNIFKGFEEFNPDRLDA 420

QY 421 DGRFRKHEAPLPSLGKRVCLGEGAKELFFFTTILQAFSLSPDPTLSLKPTVSG 480
DB 421 DGRFRKHEAPLPSLGKRVCLGEGAKELFFFTTILQAFSLSPDPTLSLKPTVSG 480

QY 481 LFNIPPAFQLOVRPTDLHSTTQTR 504
DB 481 LFNIPPAFQLOVRPTDLHSTTQTR 504

RESULT 14
ADA61287
ID ADA61287 standard; protein; 504 AA.
XX
AC ADA61287;
XX
DT 20-NOV-2003 (first entry)
XX
DE Homo sapiens.
XX
KW Human; secreted and transmembrane protein; PRO;
KW Tumour necrosis factor alpha release; TNF-alpha release;
KW Glucose uptake modulator; FFA uptake modulator;
KW cell proliferation stimulator; cell differentiation stimulator;
KW cell differentiation inhibitor; cytokine release stimulator; tumour;
KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; chromosome mapping; gene mapping;
KW gene therapy; chromosome identification; chromosome marker.
XX
OS Novel.
OS human.
OS secreted.

PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00806889.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Beresini M, Deforge L, Deanoyers L, Filvaroff E, Gao W;
PI Gerritsen WB, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Wacanabe CK, Wood WI, Zhang Z;
XX
WPI; 2003-521854/49.
DR N-PSDB; ADA18663.
XX
PT New PRO nucleic acid, useful for preparing a composition for treating
PT e.g., tumors.
XX
PS Claim 12; Fig 64; 660pp; English.
XX
CC The invention relates to isolated human PRO polypeptides (secreted and
CC transmembrane polypeptides) and the polynucleotides encoding them. The
CC invention also relates to an antibody which specifically binds to a PRO
CC polypeptide, a method for stimulating the release of tumour necrosis
CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells and a method for
CC detecting the presence of a tumour in a mammal (e.g. lung, colon, breast,
CC prostate, rectal, cervical and liver tumours). The polynucleotides are
CC useful in molecular biology, including uses as hybridisation probes, in
CC chromosome and gene mapping, in generating antisense RNA and DNA and in
CC gene therapy. The polynucleotides may also be used in preparing PRO
CC polypeptides by recombinant techniques and in generating either
CC transgenic animals or knock-out animals which are useful in the
CC development and screening of therapeutically useful reagents. The PRO
CC polypeptides or antibodies are used in preparing a medicament for
CC treating a condition responsive to the polypeptides or antibodies, such
CC as tumours, for modulating the uptake of glucose or FFA by adipocyte
CC cells, for stimulating the proliferation of or gene expression in
CC pericyte cells, for stimulating the release of proteoglycans from
CC cartilage, for stimulating the proliferation of inner ear utricular

OS and.
OS transmembrane.
OS protein.
OS PRO1906.
XX
PN US2003049816-A1.
XX
PD 13-MAR-2003.
PF
PF 15-APR-2002; 2002US-00123262.
XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019033.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 98WO-US000106.
PR 08-MAR-1999; 98WO-US005028.
PR 10-MAR-1999; 98WO-US005190.
PR 20-APR-1999; 98WO-US008615.
PR 14-MAY-1999; 98WO-US010733.
PR 02-JUN-1999; 98WO-US012252.
PR 01-SEP-1999; 98WO-US020111.
PR 08-SEP-1999; 98WO-US020594.
PR 13-SEP-1999; 98WO-US020944.
PR 15-SEP-1999; 98WO-US021090.
PR 15-SEP-1999; 98WO-US021547.
PR 03-OCT-1999; 98WO-US023089.
PR 29-NOV-1999; 98WO-US028214.
PR 30-NOV-1999; 98WO-US028313.
PR 30-NOV-1999; 98WO-US028409.
PR 01-DEC-1999; 98WO-US028301.
PR 01-DEC-1999; 98WO-US028634.
PR 02-DEC-1999; 98WO-US028551.
PR 02-DEC-1999; 98WO-US028564.
PR 16-DEC-1999; 98WO-US028565.
PR 20-DEC-1999; 98WO-US030095.
PR 20-DEC-1999; 98WO-US030911.
PR 22-DEC-1999; 98WO-US030999.
PR 30-DEC-1999; 98WO-US030720.
PR 30-DEC-1999; 98WO-US031243.
PR 05-JAN-2000; 98WO-US031274.
PR 06-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 11-FEB-2000; 2000WO-US000376.
PR 18-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
PA (GETH) GENENTECH INC.
XX
XX Baker KP, Heresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2003-695892/66.
XX N-PSDB; ADA61286.
XX
XX New PRO nucleic acid and encode polypeptides, are useful for
PT manufacturing a medicament for diagnosing or treating cancer.
PT
XX
XX Claim 12; Fig 64; 660pp; English.
XX
XX The invention describes 305 nucleic acids encoding PRO (secreted and
CC transmembrane) polypeptides (I). (I) is useful for stimulating the
CC release of TNF-alpha from human blood, for modulating the uptake of
CC glucose or FFA by skeletal muscle cells or adipocyte cells, for
CC stimulating the proliferation or differentiation of chondrocyte cells,
CC for stimulating the proliferation of or gene expression in pericyte
CC cells, for stimulating the release of proteoglycans from cartilage, for
CC stimulating the proliferation of inner ear utricular supporting cells,
CC for stimulating the proliferation of T-lymphocyte cells, for stimulating
CC the release of a cytokine from PMBC cells, for inhibiting the binding of
CC A-peptide to factor VIIA, for inhibiting the differentiation of adipocyte
CC cells, for stimulating proliferation of endothelial cells, for detecting
CC the presence of tumour in a mammal. The tumour is lung, colon, breast,
CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes
CC are useful for isolating genomic and cDNA nucleotide sequences or
CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful
CC in assays to identify other proteins or molecules involved in binding
CC interaction. A polynucleotide (II) encoding (I) is useful in chromosome
CC and gene mapping, in generation of antisense RNA and DNA, in the
CC preparation of PRO polypeptide, for generating transgenic animals or
CC knockout animals which in turn are useful in the development and

CC screening of therapeutically useful reagents, in gene therapy, for
CC chromosome identification, as chromosome marker, and for generating
CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.
CC detecting its expression in specific cells, tissues or serum, and for
CC affinity purification of PRO from recombinant cell culture or natural
CC sources. (I) and (II) are useful for tissue typing. This is the amino
CC acid sequence of a novel human secreted and transmembrane PRO
CC polypeptide.
XX
SQ Sequence 504 AA;

Query Match 100.08; Score 2615; DB 6; Length 504;
Best Local Similarity 100.08; Pred. No. 1.8e-258; Gaps 0;
Matches 504; Conservative 0; Mismatches 0; Indels 0;
QY 1 MEATGTWALLALALLLLTALSSTRARGHLPGPTPLPLGNLLQRLPGALYSGLMRL 60
Db 1 MEATGTWALLALALLLLTALSSTRARGHLPGPTPLPLGNLLQRLPGALYSGLMRL 60
QY 61 SKKYGPFVTIYLGPRPVVVLVGQAVREALGGQAEFSGRTVAMLESTPDGHGVFSN 120
Db 61 SKKYGPFVTIYLGPRPVVVLVGQAVREALGGQAEFSGRTVAMLESTPDGHGVFSN 120
QY 121 GERWQLAKFTMLARDLGMKREGELIQAARCLVETFOCTEGRPPDPSPSLLLAQATSN 180
Db 121 GERWQLAKFTMLARDLGMKREGELIQAARCLVETFOCTEGRPPDPSPSLLLAQATSN 180
QY 181 VVCSLLFLGRFSYEDKEFAVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
Db 181 VVCSLLFLGRFSYEDKEFAVRAAGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHH 240
QY 241 VSTLAFTVRQVQOQGNLDASGPARDLVDAFLKWAQEQNGFTKNNMLMTVIYLL 300
Db 241 VSTLAFTVRQVQOQGNLDASGPARDLVDAFLKWAQEQNGFTKNNMLMTVIYLL 300
QY 301 FAGTWTVTGTYTLLLMKYHVQKWREELNRELGAGQPSLGRTRLPYTDVILHEA 360
Db 301 FAGTWTVTGTYTLLLMKYHVQKWREELNRELGAGQPSLGRTRLPYTDVILHEA 360
QY 361 QRLALLVPMGIPRTLMRTFRGYTLPGQTEVPPLLSILHDPNIFKHPHEFNPDRLDA 420
Db 361 QRLALLVPMGIPRTLMRTFRGYTLPGQTEVPPLLSILHDPNIFKHPHEFNPDRLDA 420
QY 421 DGRFRKHEAFLPFLSKGRVCLGELAKAEHLFFFTTILQAFSLESPPDPTLSLKPTVSG 480
Db 421 DGRFRKHEAFLPFLSKGRVCLGELAKAEHLFFFTTILQAFSLESPPDPTLSLKPTVSG 480
QY 481 LFNIPPAFOLQVRPTDLHSTQIR 504
Db 481 LFNIPPAFOLQVRPTDLHSTQIR 504

RESULT 15
ADBL9072
ID ADBL9072 standard; protein; 504 AA.
XX
AC ADB19072;
XX
XX
DT 20-NOV-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1906.
XX
KW Human; secreted and transmembrane protein; PRO;
KW Tumour necrosis factor alpha release; TNF-alpha release;
KW glucose uptake modulator; FFA uptake modulator;
KW cell proliferation stimulator; cell differentiation stimulator;
KW cell differentiation inhibitor; cytokine release.
XX
OS Homo sapiens.
XX
XX US2003068796-A1.
XX
PD 10-APR-2003.

XX 15-APR-2002; 2002US-00123261.
PF
XX 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 14-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 06-JAN-2000; 2000WO-US000219.
PR 05-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.


```
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 01-JUN-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021056.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
XX Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2003-695927/66.
XX N-PSDB; ADB19071.
XX
XX Novel secreted and transmembrane PRO polypeptides useful for stimulating
XX the release of tumor necrosis factor alpha and detecting the presence of
XX a tumor in a mammal.
XX
XX Claim 12; Fig 64; 660pp; English.
XX
XX The invention describes 305 nucleic acids encoding PRO (secreted and
XX transmembrane) polypeptides (I). (I) is useful for stimulating the
XX release of TNF-alpha from human blood, for modulating the uptake of
XX glucose or FFA by skeletal muscle cells or adipocyt
XX
XX Sequence 504 AA;
XX
Query Match 100.0%; Score 2615; DB 6; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.8e-258;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEATGTWALLALALLLLTALSCTRARGHLPPGTPPLLCNLQLRPGALYSGLMRL 60
DB 1 MEATGTWALLALALLLLTALSCTRARGHLPPGTPPLLCNLQLRPGALYSGLMRL 60
QY 61 SKYGPVFTIYLGPRPVVVLVQEAAREALGGQAEFEFGSGRTVAMLEGTFDGHGVFFSN 120
DB 61 SKYGPVFTIYLGPRPVVVLVQEAAREALGGQAEFEFGSGRTVAMLEGTFDGHGVFFSN 120
QY 121 GERWRQLRKFTMLALDLGNGKREBELIQAEARCLVETFOGTGPRFPDPSLLLAQATSN 180
DB 121 GERWRQLRKFTMLALDLGNGKREBELIQAEARCLVETFOGTGPRFPDPSLLLAQATSN 180
QY 181 VVCSLLFGLRFSEYDEKFAVVRAGGTLGYSSQGGQTYEMFWSFLRPLPGPHKOLLHH 240
DB 181 VVCSLLFGLRFSEYDEKFAVVRAGGTLGYSSQGGQTYEMFWSFLRPLPGPHKOLLHH 240
QY 241 VSTLAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEQNPGETFTNKNMLMTVIYLL 300
DB 241 VSTLAFTVRQVQHQGNLDASGPARDLVDAFLKMAQEQNPGETFTNKNMLMTVIYLL 300
QY 301 FAGTMTVSTTVGYTLLLLMKYPHVOKWVREELNRELGAGQAPSLGDRTRLPYTDVILHEA 360
DB 301 FAGTMTVSTTVGYTLLLLMKYPHVOKWVREELNRELGAGQAPSLGDRTRLPYTDVILHEA 360
QY 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIFKHPEENPDRFLDA 420
DB 361 QRLALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLSILHDPNIFKHPEENPDRFLDA 420
QY 421 DGRFRKHEAFIPFSLGKRVCLGEGIAKAELELFFFTTILQAFSLSPCPDDTILSKPTVSG 480
DB 421 DGRFRKHEAFIPFSLGKRVCLGEGIAKAELELFFFTTILQAFSLSPCPDDTILSKPTVSG 480
QY 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504
DB 481 LFNIPPAFQLOVRPTDLHSTTTQTR 504
Search completed: September 15, 2004, 09:18:33
Job time : 134 secs
```

This Page Blank (uspto)